

Georgia State University

ScholarWorks @ Georgia State University

Applied Linguistics and English as a Second
Language Dissertations

Department of Applied Linguistics and English
as a Second Language

Summer 5-7-2010

Cognitive Factors Contributing to Chinese EFL Learners' L2 Writing Performance in Timed Essay Writing

Yanbin Lu
Georgia State University

Follow this and additional works at: https://scholarworks.gsu.edu/alesl_diss

Recommended Citation

Lu, Yanbin, "Cognitive Factors Contributing to Chinese EFL Learners' L2 Writing Performance in Timed Essay Writing." Dissertation, Georgia State University, 2010.
doi: <https://doi.org/10.57709/1396514>

This Dissertation is brought to you for free and open access by the Department of Applied Linguistics and English as a Second Language at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Applied Linguistics and English as a Second Language Dissertations by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

COGNITIVE FACTORS CONTRIBUTING TO CHINESE EFL LEARNERS'
L2 WRITING PERFORMANCE IN TIMED ESSAY WRITING

by

YANBIN LU

Under the Direction of Sara Weigle

ABSTRACT

This study investigated cognitive factors that might influence Chinese EFL learners' argumentative essay writing in English. The factors that were explored included English (L2) language proficiency, Chinese (L1) writing ability, genre knowledge, use of writing strategies, and working memory capacity in L1 and L2. Data were collected from 136 university students who received a battery of tests in two sessions. The tests consisted of timed essay writing tasks in L1 and L2, post-writing questionnaires for genre knowledge and use of strategies in the writing process, a timed grammaticality judgment task for L2 grammar knowledge, a receptive vocabulary test and a controlled-production vocabulary test for L2 vocabulary knowledge, and working memory span tasks in L1 and L2. Quantitative analyses using correlations, paired-samples *t*-test, analysis of variance and multiple regression revealed that L2 language proficiency

is the most important predictor of L2 writing, followed by genre knowledge and L2 writing strategies. L1 writing ability and working memory capacity have slight impact as explanatory variables for L2 writing performance in the timed essay writing task.

INDEX WORDS: Cognitive factors, EFL learners, L2 writing, L1 writing, L2 language proficiency, Writing strategies, Working memory capacity

COGNITIVE FACTORS CONTRIBUTING TO CHINESE EFL LEARNERS'
L2 WRITING PERFORMANCE IN TIMED ESSAY WRITING

by

YANBIN LU

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
in the College of Arts and Sciences
Georgia State University

2010

Copyright by
Yanbin Lu
2010

COGNITIVE FACTORS CONTRIBUTING TO CHINESE EFL LEARNERS'
L2 WRITING PERFORMANCE IN TIMED ESSAY WRITING

by

YANBIN LU

Committee Chair: Sara Weigle

Committee: Diane Belcher

Eric Friginal

Lucy Pickering

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2010

DEDICATION

To my family.

ACKNOWLEDGEMENTS

I am deeply indebted to a great many individuals who have helped me during my doctoral study. First of all, I would like to express my sincere gratitude to my advisor, Dr. Sara Weigle, for her guidance, support, understanding, insights, patience, and encouragement in every step of my dissertation journey. Without her and the enormous time and effort she put in, I would not have accomplished this dissertation.

I am also grateful for having an exceptional doctoral committee and wish to thank Dr. Diane Belcher, Dr. Eric Frignal, and Dr. Lucy Pickering for their comments and insights as well as their encouragement and support despite their busy schedules. My gratitude also goes to Dr. Gayle Nelson, who brought me into this wonderful community and provided continual encouragement and support all along.

I am very appreciative to the Department of Applied Linguistics for providing me with the financial support and opportunities to teach in the IEP as well as undergraduate programs. Many thanks also go to the TOEFL Grants and Awards Committee at Educational Testing Service and the Dissertation Grant Program at Georgia State University for providing grants to support the completion my dissertation project.

I would like to thank my fellow doctoral students in the Applied Linguistics program, especially Amanda Baker, Liang Guo, Jack Hardy, Eliana Hirano, Guiling Hu, Meg Montee, Kate Moran, Joseph Lee, Man Li, Jason Litzenberg, Lauren Lukkarila, Caroline Payant, Pam Pearson, Amanda Lanier Temples, Weiwei Yang, Lijuan Ye, Cheongmin Yook, and many others. I enjoyed working together with them and will remember the wonderful times in the GTA room and at our living complex Tempo Continental. My special thanks go to Liang Guo, Guiling Hu

and Man Li for their insightful discussions on my project as well as their continued encouragement, emotional support, and sustained friendship.

I would also like to extend my heartfelt thanks to my colleagues and friends in Beijing, especially Wenxia Zhang and Weimin Zhang, who helped me considerably during my data collection, including recruiting participants and essay raters and arranging facilities for testing sessions. I also wish to thank Yang Hu, Jiping Liang, Baoxia Liu, Ru Xing, Meilan Zhang, Weimin Zhang, Zhipeng Zhang for rating all the essays.

Last but not least, I wish to express my special thanks to my family, to my mother and my brother for their sincere support and encouragement all these years. I am especially indebted to my husband, Huaming Wen, for his sustained love, understanding and faith in me as well as his sacrificial support all these years, and to our daughter, Luou Wen, for being strong during the times when I was away from her and for the joy and company she brought me during her two years' stay with me in Atlanta. We have lived on three continents during the six years of my study. Now my strongest wish is for us to live together and grow old with each other.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	v
LIST OF TABLES	x
LIST OF FIGURES	xii
1 INTRODUCTION	1
1.1 Context of the Study.....	1
1.2 Research Questions	7
1.3 Significance of the Study	8
2 REVIEW OF REVELANT LITERATURE.....	11
2.1 Theoretical Frameworks in L2 Writing Research.....	11
2.2 Conceptions of Second Language Proficiency and Writing Expertise	17
2.3 Review of Relevant Studies	21
2.4 Summary	32
3 METHODOLOGIES AND MATERIALS	33
3.1 Participants	33
3.2 Instruments	34
3.2.1 Timed Essay Writing in L1 and L2.....	35
3.2.2 Writing Strategies Questionnaire	36
3.2.3 Operation Span Tasks for Working Memory Capacity.....	39
3.2.4 Vocabulary Size Tests	42
3.2.5 Timed Grammaticality Judgment Task	43
3.3 Data Collection Procedures.....	44

3.4	Scoring	46
3.5	Statistical Analysis	52
4	RESULTS.....	55
4.1	Preliminary Data Analysis	55
4.1.1	Demographic Information of the Participants.....	55
4.1.2	Descriptive Statistics and Reliability of the Variables	57
4.1.3	Inter-rater Reliability.....	61
4.2	Exploratory Analyses of the Relationships among the Variables.....	62
4.2.1	Research Question 1: Relationship between Writing Ability and Workign Memory Capacity.....	62
4.2.2	Research Question 2: Relationship between L1, L2 Writing Ability And L2 Language Knowledge	63
4.2.3	Research Question 3: Relationship between L1, L2 Writing and The Use of Strategies	69
4.2.4	Research Question 4: Relative Importance of the Selected Variables to L2 Writing Performance.....	76
5	DISCUSSION AND CONCLUSION	84
5.1	Summary of the Findings and Discussion.....	84
5.1.1	Research Question 1	84
5.1.2	Research Question 2	87
5.1.3	Research Question 3.....	94
5.1.4	Research Question 4.....	98
5.2	Implications.....	100
5.2	Limitations and Suggestions	102
5.3	Conclusion.....	105

REFERENCES	106
APPENDICES	119
Appendix A ESSAY WRITING PROMPTS	119
Appendix B QUESTIONNAIRES	120
Appendix C TASKS FOR WORKING MEMORY CAPACITY.....	129
Appendix D VOCABULARY SIZE TESTS	133
Appendix E MATERIAL FOR GRAMMATICALITY JUDGMENT TASK	144
Appendix F ESSAY RATING SCALES	146
Appendix G CORRELATIONS MATRIX.....	149
Appendix H APPROVED INFORMED CONSENT FORM	150

LIST OF TABLES

Table 3.1 <i>Structure of the Writing Strategies Questionnaires</i>	38
Table 3.2 <i>Variables and Corresponding Observations</i>	54
Table 4.1 <i>Characteristics of the Participants</i>	56
Table 4.2 <i>Self-evaluation of English Proficiency Overall, English Writing Ability, and Chinese Writing Ability</i>	57
Table 4.3 <i>Self-report of Hours per Week Spent on English and Chinese Writing</i>	57
Table 4.4 <i>Descriptive Statistics of the Observed Variables</i>	58
Table 4.5 <i>Reliability and Item Statistics</i>	59
Table 4.6 <i>Reliability and Item Statistics of Measures after Adjustment</i>	61
Table 4.7 <i>Descriptive Statistics of Measures after Adjustment</i>	61
Table 4.8 <i>Inter-rater Reliabilities for Chinese Essay Ratings</i>	62
Table 4.9 <i>Inter-rater Reliabilities for English Essay Ratings</i>	62
Table 4.10 <i>Correlations between Working Memory Capacity and Writing Scores in L1 and L2</i> 63	
Table 4.11 <i>Correlations between Content and Language Scores in L1 and L2</i>	64
Table 4.12 <i>Crosstabulation of Chinese Writing Band and English Writing Band</i>	65
Table 4.13 <i>Descriptive Statistics of English Writing Scores across Chinese Writing Bands</i>	66
Table 4.14 <i>Relationship among EK, CW and EW across EK levels</i>	67
Table 4.15 <i>Correlations between Subscales for L2 writing and L2 language knowledge</i>	68
Table 4.16 <i>Regression Summary for GR, VOC1 and VOC2 Predicting EC and EL</i>	68
Table 4.17 <i>Correlations between Subscales for Strategy Use and Writing Performance</i>	69
Table 4.18 <i>Items Showing Significant Differences in Chinese versus English Genre Knowledge and Strategy Use</i>	71

Table 4.19 <i>Frequencies and Percentages of Yes-No Answers to Item #26 & #33</i>	72
Table 4.20 <i>Items Showing Significant Differences in Chinese versus English Revision</i>	73
Table 4.21 <i>Use of L1 in L2 Essay Writing</i>	74
Table 4.22 <i>Rotated Component Matrix for Revision Items</i>	75
Table 4.23 <i>Correlations between EC / EL and Revision Strategies</i>	76
Table 4.24 <i>Descriptive Statistics of the Latent Variables</i>	78
Table 4.25 <i>Pearson Correlation Coefficients of the Latent Variables</i>	78
Table 4.26 <i>Model Summary from Stepwise Multiple Regression</i>	79
Table 4.27 <i>Model Summary from Backward Multiple Regression</i>	80
Table 4.28 <i>Coefficients from Backward Multiple Regression</i>	80

LIST OF FIGURES

Figure 2.1 Hayes-Flower (1980) writing model	12
Figure 2.2 Hayes (1996) writing model	14
Figure 2.3 Grabe and Kaplan's (1996) model of writing.....	15
Figure 2.4 Bachman & Palmer's (1996) model of CLA.....	19
Figure 4.1 Contrast of Numbers of Yes/No Answers to Difficulty and Revision	72
Figure 4.2 Normal Probability Plot of Regression Standardized Residual.....	82
Figure 4.3 Scatter Plot of Residuals.....	82

CHAPTER 1

INTRODUCTION

Context of the Study

Writing in one's native language (L1) is a complex cognitive activity. The metaphor of a writer as a busy switchboard operator trying to juggle a number of demands and constraints (Flower & Hayes, 1980) vividly describes the dynamics of a writer in action. Writing in a second/foreign language (L2) is even more demanding, more difficult, and less effective, due to the additional constraints arising from the writer's L2 proficiency (Silva, 1993). In the past decades, research into L2 writing has flourished, yet most of the studies have focused on the social/cultural or pedagogical aspects of writing. As Schoonen et al. (2009) noted, cognitively oriented research on writing is a young but growing field, and a central issue is to define a "blueprint of the writer" (p. 77), that is, what individual factors are involved in both L1 and L2 writing. This study aims at contributing to this growing area of research by exploring the relationship among various factors involved in L2 writing from a cognitive perspective.

A number of studies have investigated explanatory variables for L2 writing in various contexts (e.g., Cumming, 1989; Sasaki & Hirose, 1996; Schoonen et al., 2003, 2009; Yun, 2005), and many researchers have proposed models of writing to illustrate the factors involved in writing and their interplay in the writing process (e.g., Bereiter & Scadamalia, 1987; Hayes, 1996; Hayes & Flower, 1980 for L1 writing; Grabe & Kaplan, 1996; Sasaki, 2002 for L2 writing; also see Shaw & Weir, 2007 and Weigle, 2002 for summaries of writing models). Cognitive approaches to writing focus on the cognitive factors that are important for the individual writer in developing writing expertise. In this study, four major factors were selected for investigation—L1 writing ability, L2 language knowledge, use of writing strategies and working memory capacity.

Within the previous literature on L2 writing research, L1 writing ability and L2 language proficiency are the two most frequently investigated variables. The influence of L1 writing on a language learner's L2 writing has been a primary concern since Kaplan's (1966) seminal exploration of the different rhetorical patterns of writings produced by culturally and linguistically distinct groups of writers. Since then, a great number of studies have been carried out. In terms of second language acquisition (SLA), the influence of L1 on L2 learning has been considered negative transfer or interference; that is, the quality of L2 writing might be negatively influenced by a learner's L1 rhetoric patterns. However, according to Cummins' (1979, 1981) proposal of Interdependence Hypothesis, which postulates that there is a common underlying proficiency in terms of one's cognitive and linguistic abilities and that L1 background serves as the basis for L2 learning, one might hypothesize that a learner's L2 writing ability is dependent upon his/her L1 writing ability.

Empirical studies on the role of L1 writing ability on a learner's L2 writing have yielded controversial results. Many studies (e.g., Cumming, 1989; Schoonen et al., 2003; Sasaki & Hirose, 1996; Yun, 2005) have found evidence for the Interdependence Hypothesis, showing varying degrees of contributions of L1 writing ability to L2 writing performance. Nevertheless, contrary to those studies which found some sort of correlation between L1 and L2 writing, some other studies (e.g., Carson, Carrell, Silberstein, Kroll, & Kuehn, 1990; Pennington & So, 1993) did not find a clear relationship between L1 and L2 writing products in their investigations of L2 learners' writing performance. Both Carson et al. (1990) and Pennington and So (1993) used Chinese native speakers as their participants, which raised the question of whether such controversial findings resulted from the differences among the participants' L1 in different studies.

A learner's L2 language knowledge¹ is another explanatory variable that has been widely investigated with mixed results. For example, Cumming (1989) and Schoonen et al. (2003) reported that for their participants (French and Dutch L1, respectively) L2 writing performance was highly correlated with L1 writing ability, more than with their L2 linguistic knowledge. In contrast, Sasaki and Hirose (1996) and Yun (2005) found that for their participants (Japanese and Korean L1, respectively) L2 proficiency seemed to play a more important role than L1 writing ability in L2 writing performance. Pennington and So (1993) also claimed that L2 proficiency was a much better predictor than L1 writing ability for their Chinese L1 participants. Therefore, additional empirical research is needed in order to elucidate the respective role of L1 writing ability and L2 proficiency on a learner's L2 writing performance.

A third frequently examined factor is the use of writing strategies in the process of writing. Process-oriented studies on L2 writing have found that the quality of L2 writing is more closely related to learners' writing strategies rather than their L2 proficiency. These researchers maintain that the determining factor of L2 writing quality is not the learners' linguistic competence, but their composing competence (see Krapels, 1990 and Roca de Larios, Murphy & Marin, 2002, for critical reviews of studies on L2 writing process research). Composing competence may not be the same as writing strategies, yet the effective use of writing strategies is a good indicator of one's composing competence. Studies into the writing process and writing strategies have shown that skilled writers use different skills, behaviors and procedures from unskilled writers (Roca de Larios et al., 2002). However, most of the studies into the writing process and writing strategies used qualitative methods such as think-aloud or stimulated recall for data elicitation and investigated small numbers of participants whose results were hardly

¹ L2 language knowledge is also called "L2 linguistic knowledge", "L2 language proficiency", "L2 linguistic competence" among previous studies. These terms are used interchangeably in the current study.

generalizable. Using questionnaires to measure writing strategies has the potential to reach a larger number of participants and thus conduct studies on a larger scale than is possible through think-aloud or stimulated recall. In addition, there is a need for a catalogue which embraces the fullest possible range of writing strategies, even though the lack of such a catalogue was pointed out by Leki (1995) over a decade ago.

A fourth variable that is gaining interest in L2 writing research is working memory capacity. Working memory, as a theoretical construct in the field of cognitive psychology, refers to the storage and processing mechanism for complex cognitive activities such as reading and writing. The concept of working memory, first proposed by Baddeley and Hitch (1974), involves “the temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities” (Baddeley, 2003, p. 189). It is postulated that working memory is both a processing and storage mechanism, which consists of a central executive, responsible for the control and regulation of attention, and two subsystems, the phonological loop and visuo-spatial sketchpad, responsible for maintaining phonological and visual representations (Baddeley & Hitch, 1974; Baddeley, 1986). The coordinated functions of the three components are posited to facilitate the performance of complex cognitive tasks, with the central executive playing the most important role. According to Hayes (1996), working memory plays a central role in the activity of text generation and occupies the central position in his model of writing. Within this model, working memory serves as a maintaining and processing interface of different kinds of knowledge involved in the writing process. Such a central role of working memory is also demonstrated in Kellogg’s (1996) model of working memory in writing, who argued that all the three components of Baddeley’s (1986) working memory model were in use during the composing process of writing.

It is also posited that working memory has a limited capacity. Research in cognition as information processing suggests that cognitive resources are limited, and if processing capacity is being used for one function, other functions can only make use of whatever capacity is left over (Bereiter & Scardamalia, 1987). McCutchen (1996) proposed a capacity theory of writing on the basis of a comprehensive review of the role of working memory in the process of composition. She postulated that while composing a text, a writer coordinates within working memory the planning goals and language generation processes to retrieve words and organize them into appropriate text. That is, writing imposes considerable processing and storage demands on the writer. Due to overall resource limitations within the working memory system, fewer resources will be available for storage when more resources are devoted to processing and vice versa. Therefore, if either processing or storage functions are compromised, overall writing performance will suffer.

It seems logical to assume, then, that for language learners, when they are writing in L2, they must use part of their cognitive capacity to focus on the language so that other functions, such as higher order functions for organization and discourse, cannot be engaged at full capacity. Commenting on broader concerns of text construction, Freedman, Pringle, and Yalden (1983) asserted that “constraints of writing, without full proficiency, in a second language may impose psychological limitations on people’s abilities to conceptualize their intended meaning and its organization as discourse” (p. 10). Weigle (2005) postulated that “adults writing in their first language have automatic access to lexical and syntactic resources, while for many second language writers, particularly at lower levels of proficiency, these processes are not as automatic so writers need to focus conscious attention on them, making it difficult to access strategies and LT (long-term) knowledge that are available to them when writing in their first language” (p.

136). It is possible that more attention to language will result in less attention to more global writing functions.

Working memory capacity also differs among different individuals, and people with greater working memory capacity may be better able to handle tasks which demand complex cognitive abilities. For example, Daneman and Carpenter (1980) found that one's reading span (a measure of working memory capacity) correlates with one's reading comprehension. The usual term for the measure of working memory capacity is "working memory span" (Daneman & Carpenter, 1980). In working memory span studies, participants are typically required to carry out processing and storage simultaneously. The most typical span task is the reading span test initiated by Daneman and Carpenter (1980), in which the participants read out a set of sentences while remembering the last word in each sentence for subsequent immediate recall at the end of each set. Other measures such as listening span, speaking span, writing span, and operation span (in which the participants do mental arithmetic and remember words for recall) have also been used in empirical studies.

A number of studies with L2 learners have suggested that working memory plays a role in second language acquisition (e.g., N. Ellis, 1996, 2002; Harrington & Sawyer, 1992; Mackey, Philp, Egi, Fujii, & Tatsumi, 2002; Miyake & Friedman, 1998; Osaka & Osaka, 1992; Service, 1992; Service, Simola, Metsaenheino, & Maury, 2002; Van den Noort, Bosch, & Hugdahl, 2006; Williams, 1999). Most of the studies investigating working memory capacity and second language acquisition focus on phonological short-term memory (e.g., Mackey et al., 2002; N. Ellis, 1996), reading comprehension (e.g., Harrington & Sawyer, 1992), grammar learning (e.g., McDonald, 2006), sentence parsing (e.g., Juffs, 2004, 2005), or oral production (O'Brien, Segalowitz, Collentine & Freed, 2006). There are very few empirical studies in terms of the

relationship between working memory capacity and L2 writing ability. Therefore, part of the aims of this study is to fill in this gap by exploring the effect of working memory capacity on L2 writing performance.

In summary, research into the factors that influence an L2 learner's writing performance abounds, but results are more diverging than converging. In addition, studies in the field of L2 writing tend to focus on the variables such as L1 writing ability, L2 proficiency, and use of strategies in the writing process, whereas the role of working memory is usually investigated by researchers in the field of cognitive psychology. This study is an attempt to integrate the various fields of inquiry to explore the relationships among the various cognitive factors involved in L2 writing and to examine how much each of the factors contributes to the latent variable of L2 writing ability.

Research Questions

The purpose of this study is to combine language assessment and second language acquisition as well as research in cognitive psychology to explore some of the components of the writing expertise of L2 learners in an attempt to determine how these components contribute to the learners' L2 writing performance. The factors selected for investigation were L1 writing ability, L2 language proficiency, use of writing strategies L1 and L2, and working memory capacity. Specifically, the current study investigates the following research questions:

1. What is the relationship between Chinese EFL learners' L1 writing ability, L2 writing ability and their working memory capacity in L1 and L2?
2. What is the relationship between Chinese EFL learners' L1 writing ability, L2 writing ability and their L2 language proficiency?
3. What is the relationship between Chinese EFL learners' L1 writing ability, L2 writing ability, and their use of writing strategies in L1 and L2?
 - a. What are the similarities and differences between the use of writing strategies in L1 and L2 writing?

4. What is the relative importance of the investigated variables in the explanation of L2 writing performance? How much does each of the variables contribute to Chinese EFL learners' L2 writing performance?

Chinese EFL learners were chosen as target population for this study on account of the following reasons: first, contradictory results regarding the importance of L1 in L2 writing have been reported with Chinese L1 participants and those with other native languages; second, China has the largest population of EFL learners; however, no study has investigated the explanatory factors for Chinese EFL learners' L2 writing performance.

Significance of the Study

This study is an interdisciplinary study which connects different fields such as second language writing, language assessment, and cognitive psychology. The investigation into the selected cognitive variables of EFL learners' writing performance has a number of implications which contribute theoretically, methodologically and pedagogically to the relevant fields of applied linguistics.

Theoretically, this study contributes to a better understanding of the various factors involved in L2 writing from a cognitive perspective. Some of the factors (e.g., L2 proficiency, strategy use in L2 writing process, etc.) have been examined by L2 writing researchers whereas others (e.g., working memory capacity) have been investigated by cognitive psychologists. By integrating these factors in one study, a better, though not thorough, understanding of what makes a good L2 writer could be reached. As Grabe (2001) pointed out in his "Notes towards a theory of second language writing", there is still a lack of "a predictive model of the construct of writing that would be directly and transparently useful for research agendas, instructional practices, curricular planning, and assessment efforts" (p. 48). This study does not aim to propose a comprehensive model of writing, but does try to build an equation for predicting L2

writing performance from the measures used in this study.

Methodologically, a battery of tests was compiled and/or constructed in order to measure the various contributing factors of L2 learners' writing performance. Item analysis was conducted for these instruments, and reliability and validity of these measures were examined. The examination of a post-writing questionnaire for eliciting the writing strategies that the participants used during the writing process contributes to the construction and further improvement of such a questionnaire. The modeling of L2 writing ability as measured by the timed essay writing task adapted from TOEFL TWE writing prompts will enhance our understanding of the construct of this type of writing task, and in turn, will provide useful information for test developers as well as test users.

Pedagogically, the results from this study will have significant implications for L2 writing instruction. It would be helpful for both teachers and learners of L2 writing to have information about the contributing variables of L2 writing and their relative importance as well as the effective writing strategies used by successful L2 writers during the writing process. Teachers will be able to design their curriculum and plan their lessons accordingly. For example, if L2 language knowledge is found to be the most important predictor of L2 writing performance, teachers should highlight the priority of improving the students' L2 language proficiency in L2 writing instruction. EFL learners will benefit from learning about effective writing strategies and practice accordingly.

Despite these potential significances, it has to be pointed out that the current study focused on the cognitive aspects of writing only and investigated a group of Chinese college students who were learning English as a foreign language, so the scope of this study is limited to the particular context of Chinese EFL academic writing. It would be pompous to claim that this

study solves the puzzle of what makes a good L2 writer, but hopefully this study contributes one or two pieces to the solving of this puzzle and some strokes to the painting of the blueprint of an L2 writer.

CHAPTER 2

REVIEW OF RELEVANT LITERATURE

This chapter consists of three main sections. The first section introduces the theoretical frameworks used in L2 writing research. The second section reviews concepts of second language proficiency and writing expertise. The third section reviews studies related to the selected explanatory factors: L1 writing ability, L2 language proficiency, use of writing strategies, and working memory capacity.

Theoretical Frameworks in L2 Writing Research

It is generally acknowledged that second/foreign language writing research has developed on the basis of two distinct theoretical perspectives: L1 writing theories or models and theoretical constructs in the field of second language acquisition research related to literacy transfer (Roca De Larios et al., 2002). A number of researchers have proposed models of writing in an attempt to describe the components of the writing process and their interactions. In this section, I will review the writing models proposed by Hayes and Flower (1980) and Hayes (1996), whose work was based on studies done in L1 writing, and Grabe and Kaplan (1996), whose model took second/foreign language into consideration. I will also introduce the Interdependence Hypothesis and the Threshold Hypothesis by Cummins (1979, 1981), who proposed these two hypotheses to explain the language transfer issues among bilingual learners.

Models of Writing

Hayes and Flower (1980)

The influential model of writing that Hayes and Flower (1980) devised from their analysis of writers “thinking aloud” protocols consists of three major parts: the task environment, the writer’s long-term memory, and the writing process (see Figure 2.1). The task environment,

which includes the writing assignment and the text produced so far, and the writer's long term memory, which includes knowledge of topic, knowledge of audience, and stored writing plans, are the context in which the writing process operates. The cognitive process, which is the nucleus of the model, consists of three major components—Planning, Translating, and Reviewing. In the Planning process, the writer takes information from the task environment and the long-term memory and uses it to generate ideas, set goals, and establish a writing plan. In the Translating process, the writer transforms the information, ideas, and plans into written text (this process was later called “Text Generation” in Hayes, 1996). In the Reviewing process, the writer reads and edits what has been produced to improve the quality of the text. It should be noted that Hayes and Flower emphasized that the process described in the model is recursive rather than successive. The whole writing process, including Planning, Translating, and Reviewing, may appear as a part of an Editing sub-process. Thus, writing involves a complex intermixing of stages, which is controlled by the Monitor.

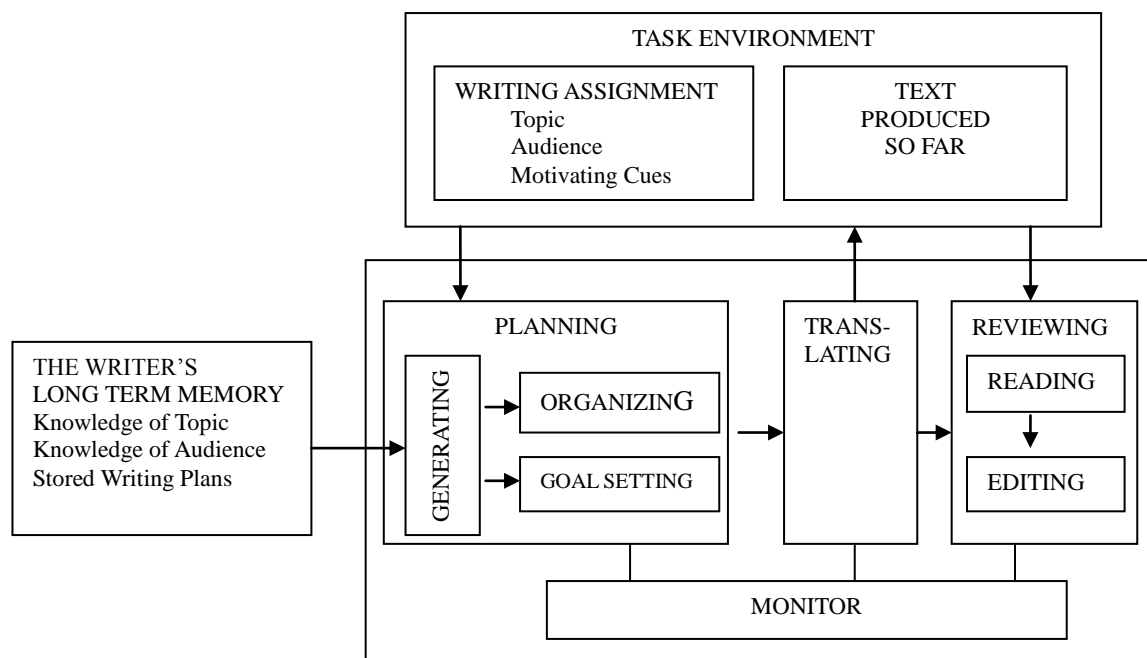


Figure 2.1 Hayes-Flower (1980) writing model

Hayes and Flower's model is important in that it served as a basis for later research into writing processes, including second language writing process research. A great many of later studies (e.g., McCutchen, 1994, 1996; Sasaki & Hirose, 1996; Zamel, 1983) were designed according to the process stages postulated in their model. In fact, this model has created the terms most commonly used by researchers in dealing with the writing process—especially the distinction between planning, formulation and revision.

Hayes (1996)

Hayes proposed a new model in 1996, which he claimed to be “a framework that can provide a better description of current empirical findings than the 1980 model” (p. 1). In this revised model, two major components—the task environment and the individual—are identified (see Figure 2.2). The cognitive processes in the new model now include text interpretation, reflection, and text production. Text Interpretation is the process by which source texts are read and internal representations are formulated. Reflection is a process of mental activities by which new internal representations are created from existing internal representations. And Text Production is the process by which the product of reflection is transcribed and elaborated into texts. These three processes are involved not only in drafting a piece of writing but also in revising one's writing as well.

Apart from other major revisions, the most important difference between the new model and the old one is that the new model emphasizes the central role of working memory in writing. Drawn heavily on Baddeley's (1986) model of working memory, the working memory in the new writing model consists of phonological memory (for phonologically coded information), visual/spatial sketchpad (for visually or spatially coded information), and semantic memory (for conceptual information). It is located in the center of the model, signifying its central role, which

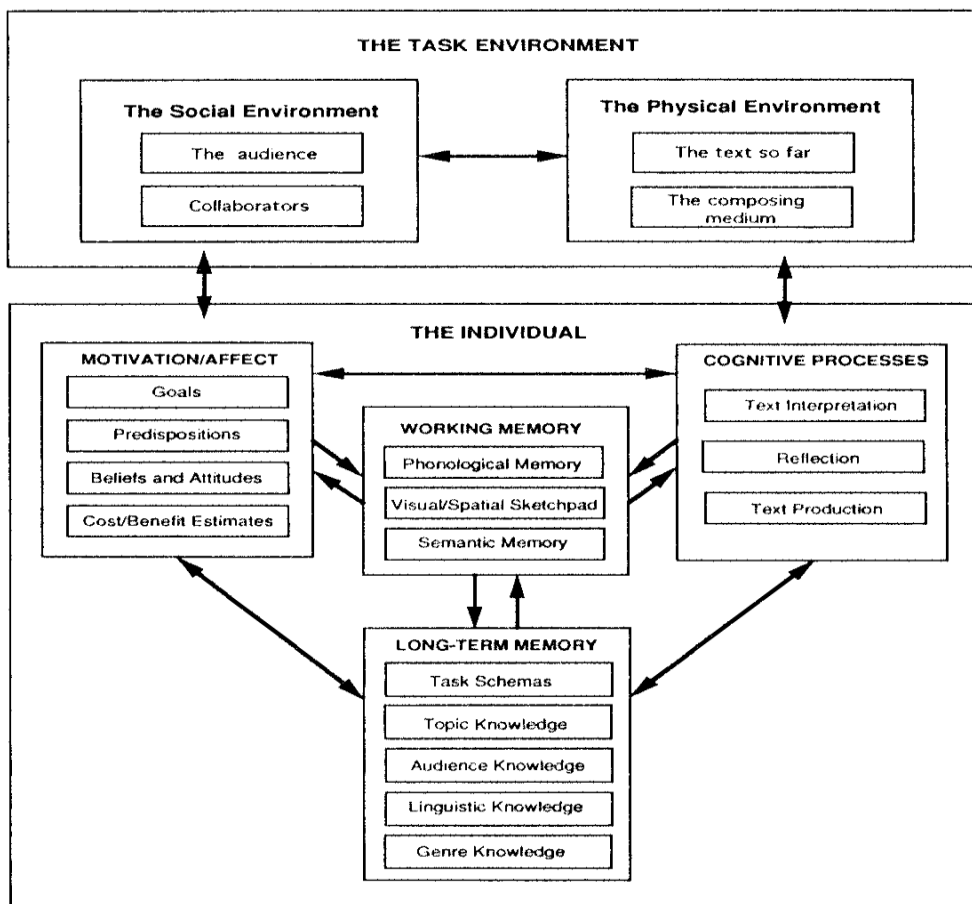


Figure 2.2 Hayes (1996) model

is to retrieve information from long-term memory and other environments and to carry out cognitive processes within its limited capacity. The role of working memory will be revisited in a late section when a review of studies on working memory and writing is presented.

Grabe and Kaplan (1996)

Attempting to integrate the cognitive processing of a writer, the linguistic and textual resources for a writing task, and the contextual factors that strongly shape the nature of writing, Grabe and Kaplan (1996) proposed a model of writing on the basis of a more general model of communicative language use developed by Chapelle, Grabe and Berns (1993). Included in the

model are two major components: *context* for language use and *verbal working memory* of the language user (see Figure 2.3). Context, as the external factor, comprises *situation* and language *performance* output. The former is further composed of participants, setting, task, text, and topic, and the latter accounts for the actual textual output produced as a result of the processing in verbal working memory.

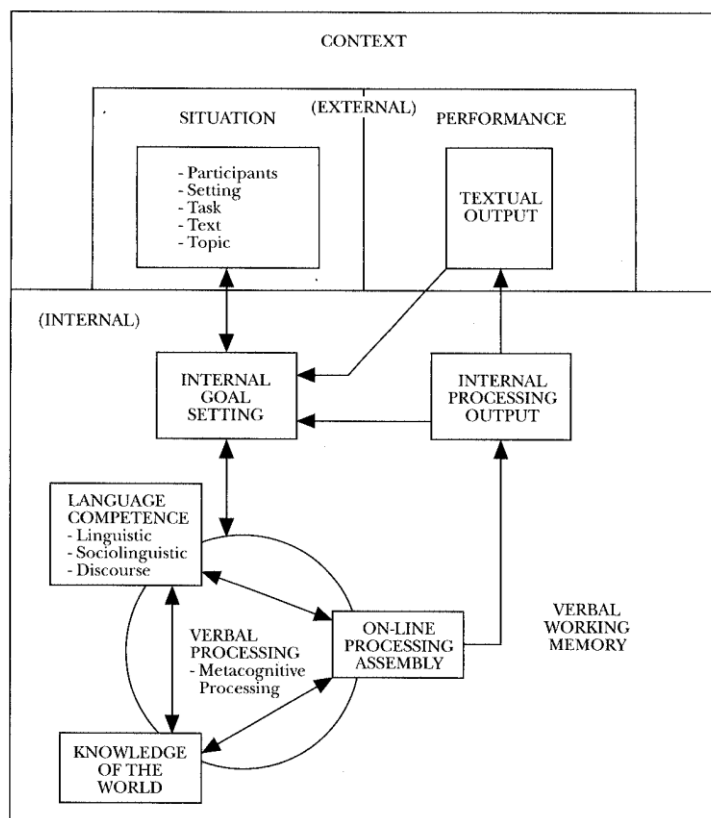


Figure 8.2 Model of writing as communicative language use

Figure 2.3 Grabe and Kaplan's (1996) model of writing

The verbal working memory is a different concept from the working memory in Hayes' (1996) model, which was built on Baddeley's (1986) conception. The verbal working memory in Grabe and Kaplan's (1996) model is constituted by three subcomponents: internal goal setting, verbal processing, and internal processing output. Internal goal setting mediates the external

context with “verbal processing,” which draws on “language competence” and “knowledge of the world” and integrates the activated resources through “on-line processing assembly.” The result of verbal processing is “internal processing output” and this processing output in turn becomes textual output as external performance. Both processing and textual output are compared to the goals set in the goal setting component which interacts with external situation and internal verbal processing.

In this model, the verbal working memory is the locus of internal operations for linguistic processing, and the relationships among the components are depicted clearly in terms of their interactions. Despite some self-acknowledged limitations of this model (such as lack of details as to how the various subcomponents of language competence interact with each other and with world knowledge and processing mechanisms), Grabe and Kaplan’s (1996) model represents an effort to incorporate social context, cognitive processing and textual product into a single model of writing from the perspective of communicative language use. Their emphasis on situation and language competence allows for a wider range of application of this model to account for not only L1 but also L2 writing.

Cummins’ Hypotheses of Language Transfer

Interdependence Hypothesis

Cummins’ notion of language transfer is expressed in two hypotheses: the Developmental Interdependence Hypothesis and the Linguistic Threshold Hypothesis. The Interdependence Hypothesis, also known as “Common Underlying Proficiency (CUP) model,” posits that “the literacy-related aspects of a bilingual’s proficiency in L1 and L2 are seen as common or interdependent across languages” (Cummins, 1981, pp. 23-24). Basically, it assumes a common set of abilities involved in cognitively demanding tasks underlying both first and second

language performance. When such a cognitive/academic proficiency is acquired in one language, it is transferrable to another language. In other words, L2 literacy skills (such as reading and writing) are partially dependent on the literacy already developed in the L1.

Threshold Hypothesis

Cummins' Threshold Hypothesis postulates that "those aspects of bilingualism which might positively influence cognitive growth are unlikely to come into effect until the child has attained a certain minimum or threshold level of competence in a second language" (1979, p. 229). In other words, in the acquisition of L2 literacy, a certain threshold or level of L2 proficiency must first be achieved if L1 literacy skills are to be transferred. Within this hypothesis, the level of a learner's L2 linguistic competence may act as an intervening variable in mediating the development of cognitive abilities and literacy skills.

It follows from these hypotheses that both L1 writing ability and L2 language proficiency are important in a learner's L2 writing development. Cummins' theory has been generated from and supported by studies involving bilingual children from a developmental perspective of their L1 and L2 proficiency. The applicability of this theory to the context of adult L2 learners merits further research.

Conceptions of Second Language Proficiency and Writing Expertise

Many researchers have explored and defined the construct of second language ability in different ways at different times. In this section, I will first review some influential models of second language ability which define the construct of language ability for the purpose of assessment and then narrow it down to components of second language writing expertise.

Early models such as those proposed by Lado (1961) and Carroll (1961) focused on competence, or rather, linguistic knowledge, instead of performance or ability to use the

language. Second language proficiency was viewed in terms of skills (i.e., reading, writing, listening, and speaking) and components (or elements) of language knowledge (e.g., grammatical structure, vocabulary, pronunciation, etc.). In contrast to the skills/components model of second language proficiency, Oller (1979) argued for an integrative notion of language proficiency (the unitary trait hypothesis). He proposed the concept of “pragmatic expectancy grammar,” which refers to the knowledge that “causes the learner to process sequences of elements in a language that conform to the normal contextual constraints of that language” (p. 38). His efforts to integrate language processing and language use highlight the importance of cognitive factors in second language proficiency. Unfortunately, as Purpura (1999) noted, he did not explicate the cognitive processes affecting second language acquisition or performance; nor did he explain how these processes could be incorporated into a comprehensive model of second language ability.

Canale and Swain (1980) proposed a model of second language proficiency which identified strategic competence as one of the three major components of communicative competence, the other two being grammatical competence and sociolinguistic competence. Later, Canale (1983) revised the model and added discourse competence to refer to cohesion and coherence. Canale and Swain were the first to include strategic competence into a model of language ability; nevertheless, their model is inadequate in that, as Chalhoub-Deville & Deville (2005) pointed out, while it is possible to view grammatical competence in terms of knowledge, it is difficult to conceptualize the other competences in their models.

On the basis of earlier work, Bachman (1990) and Bachman and Palmer (1996) formulated the influential model of communicative language ability (CLA). The CLA model represents a complex interactional framework of language use, among the various individual

characteristics of language users on the one hand, and between these characteristics and the characteristics of the language use task and setting on the other (see Figure 2.4). The key components constituting characteristics of individual language users include topical knowledge, language knowledge, personal characteristics, affective schemata, and strategic competence (those in the bold circle). Strategic competence is placed at the center of the diagram, providing links for other components within the individual as well as links with the characteristics of the language use task and setting. Bachman and Palmer (1996) defined strategic competence as “a set of metacognitive components, or strategies, which can be thought of as higher order executive processes that provide a cognitive management function in language use, as well as in other cognitive activities” (p. 70).

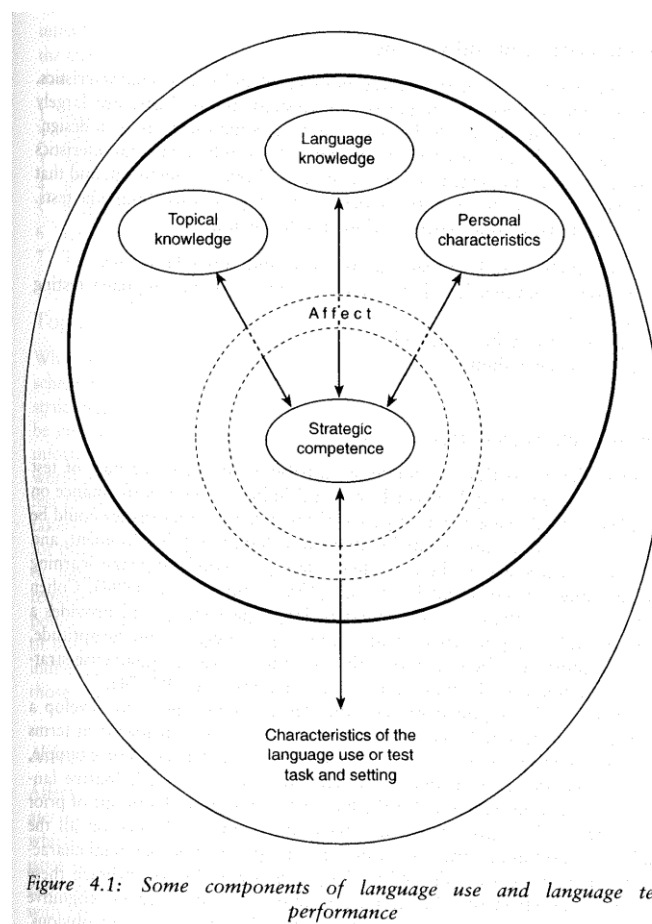


Figure 4.1: Some components of language use and language test performance

Figure 2.3 Bachman & Palmer's (1996) model of CLA

Bachman and Palmer's (1996) model presents a cognitive perspective of communicative language use. These cognitive factors would therefore apply to writing, one of the four essential language skills, which involves productive language use. In an attempt to structure the writing context and the skills, knowledge, and processes involved in writing, Grabe and Kaplan (1996) compiled a detailed taxonomy specifically for the description of the nature of academic writing. Among their twelve categories, two of them are most relevant to my current study—linguistic knowledge and writing process strategies. Linguistic knowledge includes the knowledge of the written code (orthography, spelling and punctuation), knowledge of phonology and morphology, vocabulary, syntactic/structural knowledge. Writing process strategies refer to the executive control or metacognitive strategies, among which monitoring text production, considering task problems, re-reading already produced texts and editing texts are just a few of the subcategories.

Summarizing from a cognitive perspective on writing, Weigle (2005) outlined the essential areas of knowledge that are involved in both L1 and L2 writing. These areas of knowledge include language knowledge, topic, genre, audience knowledge, task schemas and metacognition. Language knowledge includes linguistic knowledge, such as knowledge of vocabulary, grammar and orthography, as well as functional knowledge of language such as discourse knowledge and pragmatic knowledge. Weigle (2005) also pointed out the important role of working memory in writing, which will be discussed in another section later on the relationship between working memory and writing.

All in all, the review of writing models, language transfer hypotheses, and language proficiency models helped provide an understanding of the construct of L2 writing and the possible cognitive variables that contribute to L2 writing. Four aspects were selected for the current study as explanatory factors for L2 writing performance: L1 writing ability, L2 language

proficiency, use of writing strategies in the writing process and working memory capacity. In the next section, previous studies in L2 writing in relation to these four aspects are reviewed.

Review of Relevant Studies

Research on L1 and L2 Writing Ability and L2 Proficiency

Many studies have investigated a composite of variables that interplay in a learner's L2 writing performance, especially on the relationship between writing skills (both in L1 and L2) and L2 proficiency. In his seminal review, Silva (1993) examined 72 reports of empirical studies on L1 and L2 writing in an attempt to develop a clear understanding of the nature of L2 writing. More recently, Roca de Larios, Murphy, and Marin (2002) and Polio (2003) provided a critical examination of more recent empirical studies on L2 writing. Even such extensive reviews cannot conclude with a simple statement regarding the relationship between L1 and L2 writing skill and L2 proficiency, because the issues involved are complicated and the studies reviewed produced mixed results.

Empirical studies on the relationship between L1 and L2 writing and L2 language proficiency have yielded inconsistent results. Many studies (e.g., Cumming, 1989; Schoonen et al., 2003; Sasaki & Hirose, 1996; Yun, 2005) have found varying degrees of contributions of L1 writing ability to L2 writing performance. For example, Cumming (1989) found that learners' L1 writing expertise had a significant effect on the quality of L2 writing for his French participants. Sasaki & Hirose (1996) also found a significant correlation between Japanese university students' L1 writing ability and their L2 writing performance ($r = .43$). This result was corroborated in Yun's (2005) study with Korean participants, who found similar correlations between L1 and L2 writing performance ($r = .42$).

In contrast to those studies which found some sort of correlation between L1 and L2

writing, some other studies (e.g., Carson et al., 1990; Pennington & So, 1993) did not find a clear relationship between L1 and L2 writing products in their investigations of L2 learners' writing performance. In their exploratory study using multiple regression analysis, Carson et al. (1990) found that L1 writing ability was not a significant predictor for L2 writing ability for both Japanese and Chinese participants. They examined the written products of 48 Chinese and 57 Japanese students in both their L1 and L2 (English). Their results showed only a weak positive correlation ($r = .230, p < .05$) for the Japanese participants but no significant correlation ($r = -.019, n.s.$) for the Chinese participants. Another study with Chinese L1 students by Pennington and So (1993) presented similar results. In their examination of the writings produced by six Singaporean college students in both their L1 (Chinese or English) and L2 (Japanese), Pennington and So did not find a clear relationship between the quality of their L1 and L2 written products. They concluded that writing quality in L2 was correlated to learners' general proficiency level in L2 but not to the quality of the written products in L1.

L2 proficiency has also been widely investigated in L2 writing research, and the results are also mixed. Cumming (1989) reported that his French participants' L1 writing expertise (measured by written composition in their L1—French) and L2 proficiency (measured by an oral interview) accounted for large but separate portions of the variance in participants' L2 composition. L1 writing ability and L2 proficiency are found to be independent of each other. Participants with professional-level L1 writing expertise produced higher quality writing in L2 (English) than those with no such expertise, and such effect was consistent over three different types of writing (letter, argument, and summary writing). Schoonen et al. (2003) discovered with their Dutch L1 participants that L2 writing proficiency (measured by three written compositions in English) was highly correlated with L1 writing proficiency (measured by three parallel written

compositions in L1), more than with either their L2 linguistic knowledge (measured in terms of vocabulary, grammar and orthography) or the accessibility of this L2 knowledge (measured by reaction time in word retrieval and sentence building).

In contrast, Sasaki and Hirose (1996) found that Japanese participants' L1 writing ability (measured by a written composition in L1) only explained 18.3% of the variance in their L2 composition scores, whereas their L2 proficiency (measured by a comprehensive English language test for learners of English, which consisted of items on structure, listening and vocabulary) explained 52.1% of the variance, and that the unique contributions of L2 proficiency and L1 writing performance toward the variance in EFL writing performance were found to be even more remarkably different—32.6% and 1.5% respectively. Therefore, L2 proficiency seems to play a more important role than L1 writing ability in L2 writing performance. Yun (2005), adopting a structural equation modeling approach, also found that for her Korean participants, L2 proficiency was a much better predictor than L1 writing ability as an explanatory variable for L2 writing performance. Yun concluded that her data and analyses provided empirical evidence for the linguistic threshold hypothesis (Cummins, 1979, 1991), which postulates that a certain minimum level of L2 competence is required before L1 skills can transfer to L2 performance.

It is possible that participants in Yun's (2005) and Sasaki and Hirose's (1996) studies were more constrained by their L2 proficiency than those in Cumming's (1989) and Schoonen et al.'s (2003) studies. It is also possible that these controversial results were due to the language typology of participants' L1s—French and Dutch are more closely related to English, all of which belong to the Indo-European family; whereas Korean, Japanese and Chinese belong to different language families. Still another possibility is that these discrepancies arose from the various measures employed for each construct in different studies. Each study operationalized and

measured L2 proficiency in a different way—Yun with a cloze test, Sasaki and Hirose with a comprehensive English Language Test for learners of English (the structure, listening and vocabulary sections), Cumming with an oral interview and Schoonen et al. with a vocabulary test, an orthographic knowledge test and a grammatical knowledge test, thus making their results difficult to compare. Therefore, additional empirical research is required in order to explicate the relative importance of L1 writing ability and L2 language proficiency to a learner's L2 writing performance.

Research on the Use of Writing Strategies and L2 Writing

The use of writing strategies during the writing process is another frequently examined factor in L2 writing research. Despite those studies which show that learners' L2 proficiency is a major factor of learners' ESL/EFL writing performance, some early studies in second language writing process research reported that learners' writing did not seem to be influenced by their L2 linguistic proficiency (e.g., Jacobs, 1982; Raimes, 1985; Zamel, 1983). Instead, the quality of L2 writing has been found to be more closely related to learners' composing competence (see Krapels, 1990 and Roca de Larios et al., 2002 for reviews of studies on L2 writing process). Expert L2 writers with less language proficiency are not impeded in their use of global cognitive strategies in writing by their lesser ability in language; by the same token, inexperienced writers with greater fluency in English are not able to tap into more effective writing processes by virtue of their greater proficiency in English (Leki, 1996).

Research into the writing process and writing strategies of skilled/expert and unskilled/novice writers has produced mostly consistent findings. In general, expert writers are found to spend more time planning and revising their work than novice writers. They tend to work recursively, managing the whole process flexibly with a balanced interaction between

planning, generating, rereading, and revising, etc. At the discourse level, they pay more attention to the content, the organization, and the choice of words and phrases. When they revise, they tend to edit globally for content and organization rather than simply making surface local changes to the text. Skilled writers also tend to take into account the reader's expectations and handle their writing accordingly. They are also found to have a more highly developed schemata for the genres in which they are writing than less skilled writers (Weigle, 2005; also see Roca de Larios, et al, 2002 for a synthesis of studies on skilled-unskilled distinction).

Most of the process-oriented studies employed qualitative methods such as think-aloud or stimulated recall for data elicitation. Data collected through qualitative methods can yield detailed information about writing process and writing strategies; however, such data are usually collected from a small number of participants and would inevitably carry individual participants' idiosyncrasies. Using questionnaires to measure writing strategies has the potential to reach a larger number of participants and thus conduct studies on a larger scale. Petrić and Czár (2003) reported their attempt to construct and validate such a questionnaire, but they had to admit that the questionnaire they developed failed to consider the contextual factors involved in the writing process and therefore elicited relative and uninformative responses. As Leki (1995) advocated fifteen years ago, there is still a need for a catalogue which incorporates the fullest possible range of writing strategies.

Research on Working Memory and Writing

The complexity of writing has been depicted by Flower and Hayes (1980) as juggling the demands and constraints of a busy switchboard operator. Scardamalia (1981) has also described how complex the task of writing is: "Too many interdependent skills are involved, and all seem to be prerequisite to one another. To pay conscious attention to handwriting, spelling,

punctuation, word choice, syntax, textual connections, purpose, organization, clarity, rhythm, euphony, and reader characteristics would seemingly overload the information processing capacity of the best intellects” (p. 81).

In order for writers to carry on such a variety of processes simultaneously, automatization of many parts of the writing process needs to be obtained so that the processing task can be achieved with slight conscious attention (Posner & Keele, 1973, cited in Bereiter, 1980). Such performance would also require certain level of working memory capacity. In an attempt to analyze the information processing load in writing, Scardamalia (1981) examined the coordination of content schemes in expository writing among 10-14 year old children. She found that children who are presumably capable of handling three to five content schemes simultaneously in speaking tend to manage only one to three in writing. This would suggest that they are so busy with other (possibly low-order) processing needs (such as spelling out the words) that they have little attentional capacity left for higher order concerns with content.

Such a capacity theory of working memory has been supported by research such as Just and Carpenter (1992) in reading comprehension and McCutchen (1996, 2000) in writing. Based on a review of past studies, McCutchen (1996) concluded that many developmental and individual differences in writing skill could be explained with a capacity theory of writing. In complex tasks such as reading and writing, comprehension and composition impose considerable processing and storage demands. With a limited working memory capacity, novice writers are unable to deal with the complex demands imposed by the writing processes. They are constrained by the limitations and tend to depend on knowledge-telling strategies (as proposed by Bereiter & Scardamalia, 1987) and engage in non-interactive processes. In contrast, skilled writers possess fluent encoding processes for text-generation and transcription, as well as

extensive knowledge about topic, genre, and routines for coordinating writing processes.

Therefore, they are able to move beyond the limits of short-term working memory and tap the resources of long-term working memory (McCutchen, 2000).

Ransdell and Levy (1999) employed complex working memory measures to investigate the role that working memory plays in reading comprehension and text composition. They proposed a resource flexibility model, and concluded from their study that

language skills require shared working memory capacity, but differ in resource allocation priorities given to the demands of remembering, reading, and writing. In particular, writing requires resource flexibility in order for the writer to successfully attend to higher level demands such as, organization of sentences, paragraphs, and entire essays. In contrast, attending to lower-level details, such as spelling, grammar, and punctuation, are less dependent on individual differences in resource flexibility. (p. 111)

Extending this notion of working memory capacity in writing to second language writers, Weigle (2005) posited that for many second language writers, particularly those with lower L2 proficiency, their access to L2 lexical and syntactic resources is not yet automatic, so they need to pay conscious attention to these processes. With a limited capacity of working memory, this consumption of attention in turn hinders their access to higher level strategies and knowledge bases that might be available to them when writing in their first language (p. 136).

In the field of second language acquisition, many of the studies investigating the role of working memory capacity focus on phonological short-term memory (PSTM) (e.g., Mackey et al., 2002; N. Ellis, 1996), reading comprehension (e.g., Harrington & Sawyer, 1992; Miyake & Friedman, 1998), grammar learning (e.g., McDonald, 2006), sentence parsing (e.g., Juffs, 2004, 2005), or oral production (O'Brien et al., 2006). Various measures for working memory capacity have been employed (see discussion on these measures in the following section) and results tend to converge on the points that L2 learners generally have lower working memory spans in their

L2 than in their L1 and that L2 span generally correlates with L2 comprehension (but see Juffs 2004, 2005 for contradictory findings with sentence parsing).

Working memory span has proved to be a good predictor of a wide range of complex cognitive skills. Harrington and Sawyer (1992) administered reading span tasks to native Japanese learners of English in both L1 and L2. Their results showed that both span measures correlated with each other and that those readers with higher L2 reading span did better on the L2 reading comprehension tasks as measured by TOEFL grammar and reading sections. Similar correlations were also reported by Miyake and Friedman (1998), who tested L2 sentence interpretation with Japanese learners of English. Mackey et al. (2002) examined the relationship between working memory capacity in L1 and L2 and the noticing of interactional feedback. Their analyses yielded high positive correlations between L1 and L2 working memory scores and marginally significant relationship between working memory scores and noticing.

McCutchen and her colleagues conducted empirical studies to examine developmental and individual differences in working memory capacity, linguistic fluency, and writing skill in both children and adults (McCutchen, 1994; McCutchen, Covill, Hoyne, & Mildes, 1994). Their results showed that for both age groups, large working memory spans were associated with better writing. In a study to determine the type of working memory measures that predict individual differences in writing performance, Randsdell and Levy (1999) used reading, speaking, and writing span tests to measure participants' working memory capacity. Their experimental results showed that all the span measures can predict both reading comprehension and writing performance.

The working memory research in SLA shows a major interest in reading comprehension, sentence processing, and oral production but a neglect of writing performance, even though it is

widely accepted that writing involves complex cognitive efforts. The few studies that did focus on the relationship between working memory capacity and L2 writing used measures for working memory capacity (such as reading, speaking or writing span) that are correlated with L2 language proficiency. In the next section, measures for working memory capacity that are commonly used in research will be reviewed.

Measures of Working Memory Capacity

In terms of methodology, a number of tasks have been used to measure one's working memory capacity, which originated from research in cognitive psychology. Both simple span tasks (such as digit and word span) and complex span tasks (such as listening and reading span) have been developed. Simple span tasks involve mere storing and recall of strings of digits or words and are believed to tap into the more traditional concept of short-term memory for storage. Complex span tasks involve both processing of auditory or visual stimuli and storing and recall of digits or words and are believed to be able to tap into both storage and processing capacity of working memory.

There are many types and variations of complex span tasks, such as backward digit span (recalling strings of digits in a reverse order as presented), backward word span (recalling strings of words in a reverse order as presented), alpha span (recalling target words in alphabetical order) (see Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005 for a more elaborated overview of working memory span tasks). According to Gass and Mackey (2007), three commonly used types of complex span tasks are operation span, counting span, and sentence span. Operation span tasks (Turner & Engle, 1989) require participants to verify arithmetic operations (sometimes correct and sometimes not) and try to remember words presented after each equation. Participants need to recall the words after reading through a set of operations

(generally two through five equations). Counting span tasks (Case, Kurland, & Goldberg, 1982) require participants to count shapes and remember the total number of shapes for later recall. Sentence span tasks (Daneman & Carpenter, 1980) involve comprehension of sentences and recall of words (either the last word of each sentence or an unrelated word presented after each sentence). Sentences can be presented as either auditory (listening span) or visual (reading span) stimuli.

One of the most widely used working memory tasks is the Reading Span Test (a specific type of sentence span tasks), developed by Daneman and Carpenter (1980). It was designed to tap both the storage and processing functions of working memory. In the original version of reading span test, participants were instructed to read aloud sets of sentences while trying to remember the final word of each sentence. The sentences were presented on cards, one sentence on each card. At the end of a set (ranging from two to five or six), a blank card would appear, signaling the participant's turn to recall the final words seen in that set in correct order. The task contained three sets each of two to six sentences. The participants began with the two-sentence set and worked up to the more difficult ones (those with more sentences). The task was terminated when the participants failed all three sets at a particular level. The level at which the participant was correct on two out of three sets was taken as the measure of his/her reading span.

Later versions of reading span tests introduced acceptability judgment of sentences (Turner & Engle, 1989; Waters, Caplan, & Hildebrandt, 1987, cited in Waters & Caplan, 1996), recalling of an unrelated word or letter following each sentence (e.g., Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Hambrick, Tuholski, Wilhelm, Payne, & Engle, 2004), randomization of presentation order (Engle, Cantor, & Carullo, 1992), and different scoring methods of the measure (see Conway et al., 2005; Friedman & Miyake, 2005). In general, reading span tasks

have been found to be a reliable and valid measure of working memory capacity.

Stimulated by the question whether working memory capacity is task dependent, that is, whether different measures/tasks yield different results in one's working memory capacity, Engle and his colleagues developed and validated operation span tasks for measuring working memory capacity. The first operation span task (Turner and Engle, 1989) used mathematical operations to replace the sentences in Daneman and Carpenter's (1980) task while retaining the task demands (processing some information and remembering and recalling a set of target words). More specifically, in the operation span task, participants are required to read aloud a mathematical equation and verify if the equation is correct or not (e.g., $(9/3)-2=1$ or $(9/3)-2=3$). Immediately after the participant says Yes or No to the equation, a word is shown. The participant is instructed to read aloud the word and try to memorize it for later recall. Then another mathematical equation is presented and the participant repeats the process until the end of a set (ranging from two to five or six equations and words) when the participant is prompted to recall all the words in that set.

Turner and Engle's (1989) experiments showed that working memory capacity measured with the operation span task correlated highly with scores on reading comprehension as well as the working memory capacity measured with the reading span task. Regression results suggested that both spans were similar in their ability to predict reading comprehension. Analyses with stepwise multiple regression even demonstrated a larger contribution of operation span than reading span to the prediction of reading comprehension (Entering operation span first accounted for 16.03% of the variance while reading span added 3.26%. Entering reading span first accounted for 13.42% of the variance while operation span added 5.87%). Therefore, it may be concluded that operation span is as powerful a predictor as reading span of one's reading

comprehension. More importantly, working memory capacity seems to transcend task, and processing differences may result from individual differences in working memory capacity.

Summary

In this chapter, writing models proposed by Hayes and Flower (1980), Hayes (1996), and Grabe and Kaplan (1996) as well as the Interdependence Hypothesis and the Threshold Hypothesis proposed by Cummins (1979, 1981) were introduced as theoretical frameworks for the current study. Models of second language proficiency and writing expertise were also reviewed to provide further understanding of the construct of L2 writing. As a result, four aspects were selected as explanatory variables for L2 writing performance for the current study. The review of the previous studies in L2 writing in relation to these four aspects provided further rationale for their selection as explanatory variables for investigation as well as a foundation for the selection of instruments for data collection.

CHAPTER 3

METHODOLOGIES AND MATERIALS

In order to investigate the research questions, a battery of tests was used to measure a number of Chinese EFL learners' writing ability in both L1 (Chinese) and L2 (English), their working memory capacity in L1 and L2, vocabulary size in L2 (both receptive and productive), grammar knowledge in L2, and the use of writing strategies in the process of writing (as measured by post-writing questionnaire). Specifically, the actual procedure of data collection for the study is spelled out in the following sections.

Participants

The data for this study were collected from students enrolled at several universities in Beijing, China in the fall semester, 2009. The target population for this study is adult Chinese EFL learners of academic English. Chinese EFL learners were targeted as participants on the account that a) some studies with Chinese L1 participants showed contrary results to other studies with different native languages, b) most studies in this field were conducted in ESL settings rather than EFL settings, and c) China has the largest population of EFL learners. College students were considered for sampling because a lot of literature investigated this group of people (as opposed to middle school or high school students) and academic writing is important for college study. In addition, they are adult, learning English for academic purposes; they had learned EFL for at least six years in the secondary school and obtained a certain degree of proficiency (they all passed the College Entrance Examinations including the subject of English); and they are expected to have the motivation to improve their English (they are expected to pass an exit exam of English proficiency test, which includes essay writing as an integral part).

I contacted one of the deputy deans of the Department of Foreign Languages at Tsinghua University, where I used to work as an English teacher, and got permission to recruit participants and conduct my study there. The participants were recruited through the university bulletin board system (BBS), posted flyers on campus, as well as announcements of several English teachers in their classes. In the recruitment flyer, a detailed description was given in regard to what would be expected of the participant. It was made clear that they were to take a series of tests in two sessions—first with paper and pen and then on computer, that their participation in the study would be totally voluntary, and that their participation would not have an impact on their English class evaluations. They were also told that they would be paid 100 RMB (approximately \$15) for a total of three hours' participation.

A total of 136 students (both undergraduate and graduate) participated in this study (demographic information will be presented in the next chapter). They were solicited under the approval of Institutional Review Board (IRB) at Georgia State University (GSU) (IRB # H10119. See Appendix H for the approved Informed Consent Form). The majority of the participants are matriculated students at Tsinghua University, with four participants coming from three other universities—they saw the recruitment posting on the BBS and contacted me to volunteer for participation. All the participants took a series of tests in two sessions for a total of about three hours (the actual data collection procedure will be reported in a later section in this chapter). At the end of the second session, each participant was paid, as promised, 100RMB in cash for their time and effort.

Instruments

The data for this study were collected using a battery of tests, which included a) timed essay writing in L1 (Chinese) and L2 (English), b) a writing strategies questionnaire after each

essay writing task, c) vocabulary tests (including receptive and controlled-production tests), d) a timed grammaticality judgment task, and e) working memory span tasks in L1 and L2. This section provides a detailed description of the instruments employed in this study, the rationale for choosing these measures, and the development of these instruments.

Timed Essay Writing in L1 and L2

Participants were given 30 minutes each to write an essay in L1 and another essay in L2 on a certain prompt. The prompts were adapted from TOEFL Test of Written English (TWE) writing prompts (ETS, n.d.). Several prompts were devised and comments and suggestions were collected from professors in my committee, fellow doctoral students at GSU, and English teachers at Tsinghua University. The following two prompts were used in the final data collection:

Prompt A: Some people choose their major field of study based on their personal interests, while others are more concerned about future employment possibilities. What position do you support? Use specific reasons and examples to support your answer.

Prompt B: Which factor do you consider most important in your decision when applying for a college/university, the major you are interested in, the reputation of the college/university, or any other factor that is more important for you? Use specific reasons and examples to support your opinion.

The rationale for choosing these prompts was as follows:

First, the mode of timed essay writing was adopted because essay writing is the most efficient and reliable way to assess writing ability. It has to be acknowledged that timed impromptu writing test has its limitations, for example, one's writing ability is judged based on a single sample which is written in a limited time frame on a given prompt (e.g., Brown, 2004; Shaw & Weir, 2007; Weigle, 2002). However, the use of timed impromptu writing tasks is a common practice in large-scale English proficiency tests, such as TOEFL TWE and iBT independent writing for international students who intend to apply for American universities, College English Test (CET, Band-4 and Band-6) for non-English major college students in China,

and Tsinghua English Proficiency Test for matriculated students at Tsinghua University. The time allowance of 30 minutes was decided on the basis of common practice as well as the research done by Jacobs et al. (1991) which found that this amount of time probably gave most students enough time to produce an adequate sample of their writing ability.

Second, the genre of argumentative prose was employed because it is believed that the ability to generate and organize ideas with examples or evidence for this type of writing involves complex cognitive functions (Hale et al., 1996). Moreover, argumentative essay writing is one of the common essay genres college students may encounter across the curriculum. It is also typical of large-scale writing tests.

Third, the two specific prompts were chosen because the topics were believed to be closely related to the participants' life as a college student; therefore, it was hoped that the participants would find it relatively easy to write about these topics (This was confirmed by informal chats with some of the participants after the written test). Specifically, the prompts required the participants to write about the most important factor(s) they considered while choosing their major field of study (for English essay) and while deciding on their target college/university (for Chinese essay). They had thought these questions over in their real lives (just recently for freshmen and a little earlier for other students); therefore, these topics were general, familiar topics that were relevant, potentially interesting, and accessible to all the participants.

Writing Strategies Questionnaire

Both the L1 and L2 writing tasks were immediately followed by a questionnaire designed to elicit participants' use of writing strategies during the process of writing the Chinese and the English essays respectively. Writing strategies in this study are defined as actions or behaviors

carried out by participants in completing the timed essay writing task as reported in the post-writing questionnaire. That is to say, this study focuses on participants' perceptions of the writing strategies they used in the writing process, rather than direct observations of what they did in the writing process. The questionnaire asked about what the participants did before, during, and after the writing of the essays. It also included items which asked about the difficulty and familiarity of the task, the genre and the topic. Therefore, the use of the post-writing questionnaires aimed at tapping into participants' self-reported genre knowledge and strategy use in the process of their essay writing.

Both statements and open-ended questions were included in the questionnaires. The statements were adapted from previous studies on writing which used questionnaires on similar issues. The major sources included Oxford (1990), whose Strategy Inventory for Language Learning (SILL) contains strategies for writing; Petrić and Czár (2003), who validated a writing strategies questionnaire; Sasaki & Hirose (1996), who used questionnaires on writing processes for both L1 and L2 essay writing; and Weir (2005), who developed a Cognitive Processing Questionnaire which reflects a theory-based framework of writing.

The questionnaire was constructed in English and translated into Chinese, though the final versions for both questionnaires were written in Chinese in order to facilitate maximum comprehension and avoid misunderstanding of the items. The English version was checked for content validity by obtaining expert opinions from two professors on the relevance of items, wording appropriateness and possible interpretation problems, as well as clarity of instructions. The Chinese version was piloted with two fellow doctoral students at GSU who are native speakers of Chinese. Wording and interpretation problems were discussed and additional ideas were elicited. As a result, some major changes were made, including eliminating irrelevant items,

adding new items and the choice of “not applicable,” reordering the items, and addressing a number of wording problems.

The final version of the questionnaire for Chinese essay writing consisted of 49 items and the one for English essay writing consists of 54 items, with five more items on the use of L1 in the writing of L2 (See Appendix B for the English version of the questionnaire). The items were sequenced following the structure of the writing process according to Hayes and Flower’s (1980) model, i.e., planning, composing and revising, so as to provide the participants with a clear frame of reference. Conceptually, the final questionnaire for Chinese essay writing consisted of five sections and that for English essay writing six sections (see Table 3.1 below), though none of the section numbers or section titles appeared in the actual questionnaire.

Table 3.1 *Structure of the Writing Strategies Questionnaires*

Section	Section Title	Number of Items
1	Genre and topic knowledge	9
2	Planning before writing	7
3	Composing during writing	9
4	Difficulties during writing	6 (7 for English essay)
5	Revision during and/or after writing	16
6 (only for English essay)	Use of Chinese	4

A yes-no question preceded Section Four and Section Five, respectively, asking the participants if they had difficulties during the process of writing and if they reread and revised their essay during and/or after the writing. If their answer to either question was “No,” they were instructed to skip the relevant section and proceed to the next section.

Five-point Likert scale was used for all the statements, following Oxford’s (1990) SILL. Participants were instructed to indicate the degree of how true each of the statements was in terms of what they actually did during their essay writing process. The scale ranged from 1 (strongly disagree) to 5 (strongly agree). A choice of “not applicable” (N/A, represented by

number 0) was given at the end of each statement. The participants were instructed to choose this scale only if the situation described by the statement never occurred to him/her.

In addition to the statement items, four open-ended questions were included out of a concern that the statements in the questionnaire were not an exhaustive inventory of what the students might have done during the whole writing process. Three questions asked for what else was done before writing, during writing, and for revising, respectively (placed at the end of the corresponding section), and one question asked for additional comments regarding the questionnaire itself or the study as a whole. Participants were instructed to answer the open-ended questions in either Chinese or English, whichever language they felt more comfortable with.

Operation Span Task for Working Memory Capacity

Operation Span Tasks were used as a measure of participants' working memory capacity in L1 (Chinese) and L2 (English). The operation span task for English was modified from the one developed by Unsworth, Heitz, Schrock and Engle (2005). It has been used as an established test for working memory capacity in the field of psychology. The operations used in the task are simple arithmetic operations involving addition, subtraction, multiplication and division (e.g., $(10 \times 1) - 7 = 3$). The English words for recall are all high frequency words with only one syllable and four to six letters. To measure working memory capacity in L1, I compiled a list of Chinese words and consulted native Chinese speakers to make sure that they are high frequency words in Chinese. See Appendix C for the specific items.

The operation span task was chosen over the reading span task on the basis of the following reasons:

First, a number of studies (e.g., Service et al., 2002; van den Noort et al., 2006) have

shown that working memory capacity as measured by the reading span task interacts with language proficiency. According to Service et al, (2002), significant differences in reading span were reported between native speaker group versus foreign language group as well as more proficient versus less proficient foreign language groups. The results of Van den Noort et al.'s (2006) study confirmed that participants' working memory capacity as measured by reading span tasks was highly correlated with their language proficiency. In order to avoid the confounding relationship between reading span and language proficiency, the reading span task was not employed.

Second, the operation span task was chosen because, according to Turner and Engle (1989) and La Pointe and Engle (1990), working memory capacity is independent of tasks and operation span tasks can measure working memory capacity as equally well as reading span tasks and serve as a powerful predictor of reading comprehension. Unsworth et al. (2005) demonstrated that the operation span task correlated well with other measures of working memory capacity (including reading span task) and that it had good internal consistency and test-retest reliability. They claimed that the operation span task was a reliable and valid indicator of working memory capacity. The reliability and validity of the operation span task have also been attested in the comprehensive review of working memory span tasks by Conway et al. (2005).

In addition, my small-scale trial of both reading span tasks and operation span tasks on some Chinese learners of English cautioned me that it would be more difficult for my potential participants to handle reading span tasks than operation span tasks due to the complexity involved in reading span tasks and their English language proficiency. My trial participants reported that during the experiment, the target words (the words that were to be memorized and recalled) would mix with the other words in the sentences, thus making the task very demanding.

They found it frustrating because they could not perform the task well. In comparison, operation span tasks were reported to be easier to handle and perform.

In the experiment, the operation span task was presented on a computer with DMDX (version 3.3.1.1), the software for visual display that was developed by John Forster and was downloadable for free (Forster & Forster, 2003). During the task, a target word was first shown on the computer screen for half a second, followed by an equation involving simple arithmetic operation (e.g., IS $(10 \times 1) - 7 = 3$ CORRECT?). The participants were instructed to perform two tasks: one was to remember the target word for later recall, and the other was to judge as fast as possible whether the equation was correct or not by pressing a corresponding key on the keyboard that was marked "CORRECT" (the right shift key) or "INCORRECT" (the left shift key). As soon as the judgment was made, another target word would appear on the screen. The participants were instructed to repeat the process (reading and memorizing the target word and judging the correctness of the following equation) until they saw a prompt on the screen which said "RECALL and write down the target words within this set." At this prompt, they should write down, on a separate answer sheet, as many target words as they could recall from this set. There was no time limit for the recall. After recalling, they pressed a key marked "NEXT" (the space key) to proceed to the next set. There were a total of 15 sets with a total of 60 items. The number of items varied within each set, with two items as the minimum and six as the maximum. In other words, there might be two target words to recall and two arithmetic operations to judge within a set (called two-item set), or there might be three, four, five, or six items. It is self-evident that the larger the number of items within a set, the more demanding it would be on one's cognitive function, that is to say, the larger the working memory load the task exerts, the harder it would be for the participant to recall all the target words.

During the actual data collection for WM capacity, the participants were, first of all, given instructions in Chinese about what to do for this task. Then they were given three sets of practice items to get themselves familiar with how the task actually proceeded. Questions were answered and mistakes corrected (e.g., some participants started writing down the target words before the “RECALL” sign appeared) before the actual test items were shown. At the end of the practice, they were told to proceed to the actual test items when they felt ready. On the whole, the participants found the tasks both manageable and interesting. Most of them reported after completing the experiments that the task for recalling Chinese words (L1) was much easier than that for English words (L2). In effect, some of the participants performed very well, with perfect or near perfect scores in the task for Chinese (recalling all or almost all of the 60 words without an error); therefore there might be a ceiling effect, which means that the majority of the scores for working memory capacity in Chinese were at or near the maximum possible for the test. Such a ceiling effect would have underestimated the working memory capacity of these participants and would cause a reduction in the measured variance and hence compromise the truthfulness of the measurement and the corresponding statistical analysis.

Vocabulary Size Tests

Vocabulary is an indispensable component in any model of language competence, and vocabulary size is generally believed to be a good indicator of a learner’s linguistic knowledge. As Read (2000) has pointed out, adequate knowledge of vocabulary is a prerequisite for effective language use. Receptive vocabulary knowledge is important for comprehension such as listening and reading, whereas productive vocabulary is important for language production such as speaking and writing. In the current study, both receptive and productive vocabulary knowledge were tested. Two vocabulary size tests developed by Nation and his colleagues were used to

measure participants' vocabulary size. The test in Nation and Gu (2007) was used for receptive vocabulary knowledge, while the test developed by Laufer and Nation (1999) was used for productive ability in vocabulary use (see Appendix D for the two tests).

The receptive vocabulary test consisted of 13 sets of questions, with ten multiple choice items within each set (a total of 130 items). The first ten items test the first thousand word level, and the second ten items test the second thousand word level, and so on, and so forth, with increasing difficulty for later items. For each item, the target word is given first, followed by a sentence with the target word in brackets. Then four choices were listed. The participants were supposed to choose the best meaning for each word. The test was printed on a 10-page double-sided test booklet. The participants were provided with a separate answer sheet in addition to the test booklet. They were instructed to record all their answers on to the answer sheet (for the sake of data storage as well as easy scoring).

The controlled-production vocabulary test consisted of five sets of 18 incomplete sentences (a total of 90 items). There was a blank in each sentence. On the blank, the first two to five letters of the expected answer was given. The participants were supposed to complete each sentence with an appropriate word starting with the given letters. The test was printed on three pages and a separate answer sheet was also attached to the test booklet for the participants to record their answers on.

Timed Grammaticality Judgment Task

A timed grammaticality judgment task (GJT) developed by Rod Ellis (2005) was used as a measure of participants' implicit grammatical knowledge (see Appendix E for all the items used in the task). Implicit, or automatized, grammatical knowledge was aimed at because it is believed that with automatized grammatical knowledge, one's processing of the language may

take up fewer resources in one's working memory, thus leaving more resources for other functions or resulting in more efficient processing of the language. GJTs are popular tasks in the field of SLA as a measure of L2 learners' linguistic abilities, especially in the investigation of acquisition of grammatical structures. According to Ellis (2005), the timed GJT is a valid measure to tap into participants' implicit knowledge of grammar.

In this timed grammaticality judgment task, the participants were required to read an English sentence and judge whether it is correct in terms of grammar by pressing one of the two keys marked "CORRECT" and "INCORRECT." They were instructed to react (both read and judge) as quickly as possible because the sentence would only be shown on the screen for a few seconds (four seconds to be exact). If they didn't make the judgment within the time limit, the sentence would disappear automatically and the next sentence would follow, and the software program would record this item as "No response." There were a total of 68 items, half grammatical and the other half ungrammatical. Each sentence contained 12-16 words and appeared on the center of the computer screen in one line for a limited period of time. There were six practice sentences at the beginning of the task for the participants to get familiar with the whole process.

Data Collection Procedures

The main data collection was conducted during the fall semester of 2009 at Tsinghua University. Participants took a series of tests in two separate sessions. In the first session they met in a classroom on a designated date and took written tests with paper and pen. The group test was administered on four different nights in order to accommodate participants' schedules. The number of participants ranged from 24 to 43 for each group test. In each of the administration, I was the only proctor they interacted with. Each group test administration lasted for 130 to 140

minutes.² The procedure was as follows:

1. Brief introduction of myself and my project (2 minutes)
2. Consent form & background information sheet (5 minutes) (during this time, I passed out a slip with a research ID to each of the participants)
3. Essay writing in English/Chinese (30 minutes) + post-writing questionnaire (10 minutes)
4. Receptive/productive vocabulary test (20-25 minutes)
5. Essay writing in Chinese/English (30 minutes) + post-writing questionnaire (5-10 minutes)
6. Productive/receptive vocabulary test (20-25 minutes)

Two groups did the English essay writing and the receptive vocabulary test first whereas the other two groups did the Chinese essay writing and the productive vocabulary test first. The participants were given test papers or booklets one test at a time, so that the time allotment for each test was guaranteed and the participants would not rush through the tests (the drawback was that those who were fast finished ahead of time and had to wait whereas those who were slow could not finish). At the end of the written tests, the participants signed up for the second session of the project—the computer tests.

In the second session, the participants took computer tests individually on a computer. Five computers were set up in a room for the computer tests. The participants came at their signed-up time slots to take the tests. The computer tests included three tasks and lasted for 40-50 minutes: a) timed grammaticality judgment task (approximately 10 minutes), b) working memory span task in Chinese (about 15-20 minutes), and c) working memory span task in English (about 15-20 minutes). Before each task, the participants were provided with written instructions in Chinese about what to do. They read through the instructions and asked any questions they had. They were also given verbal instructions which highlighted the requirements of each task. Then they read the instructions once again on the computer screen, followed by practice items and then the actual test items. At the end of each task, each individual participant

² There might be a fatigue effect on the participants due to the long hours of testing.

would let me know that he/she had come to the end. I would save the data for each of them and open up the next task. After the completion of all the three tasks, each participant was paid 100RMB for their time and efforts.

Measures have been taken to make sure that all the data were as complete and valid as possible. For example, at the beginning of the written test, I called on the participants to take the tests seriously and do their best to demonstrate their proficiency.³ In addition, I tried to obtain complete data from each of the participants. During the written tests, I browsed through the previously collected test papers and questionnaires while the participants were working on a later test so as to make sure that all the questions had been answered. If there was a question or an item left unanswered, I would ask the participant to complete it during the break. During the computer tests, I saved the data for each participant on each task. Therefore, almost all the data were complete and valid for analysis.

Scoring

Timed Essay Writing in L1 and L2

An analytic rating scale was used for the essay rating. Analytic scoring is believed to be able to provide more detailed information about a test taker's performance in different aspects of writing than holistic scoring. It is more suitable than holistic scoring especially for second language learners who may have a disproportionate development in their L2 writing ability—they might be better at content and organization but poorer in terms of vocabulary and language use (see Weigle, 2002, for a comprehensive discussion on different types of rating scales).

For my project, the rating scales and scoring rubrics for English essays were developed on the basis of several sources, including TOEFL iBT independent writing rubrics (ETS, 2008),

³ In fact, some participants asked who sponsored me for the remuneration of their participation I told them that I was paying them out of my own resources (at that time I had not been awarded any grant yet) and hoped that I would be able to collect authentic data from them.

Jacobs et al.'s (1981) scoring profile, and Tsinghua English Proficiency Test (TEPT-I) scoring guide for writing (Tsinghua University, 2009). The rubrics for Chinese essays were developed on the basis of the scoring guides for composition for the subject of Chinese Language Arts in the College Entrance Examinations (Ministry of Education, 2009). Two categories were included in the rubrics: Content & Organization and Language Use. Originally a rating scale of 1-6 bands was adopted. In the actual rating process, both the English and Chinese raters found it hard to differentiate the essays with band scores. Considering that the raters are more familiar and experienced with a more detailed scale (1-15 points for English essay rating and 1-100 points for Chinese essay rating), the rating scales were converted into point scores (see Appendix F for the rating scales and scoring rubrics).

A total of seven raters (six main raters and one third rater) were involved in the essay rating. Their ages ranged from 35 to 48 and their years of teaching experiences ranged from 10 to 26. Two of the raters were male and the other five were female. Six of them rated the 136 essays (either English or Chinese) and one (the third rater) rated 27 English essays. Four of them were paid for their time and effort while the other three declined the remunerations in exchange for a free meal and sustained friendship.

Initially, each essay was to be rated by two raters respectively. Two Chinese teachers from the Department of Chinese rated the Chinese essays and two English teachers from the Department of Foreign Languages were recruited to rate the English essays. The Chinese essays and the English essays were rated in separate sessions.

The English essays were rated first. At the beginning of the rating session, the two raters met together with me in an office. I briefly explained the purpose of the rating and the two categories as well as the descriptors in the scoring guide. The two raters read through the rating

scales (with band scores of 1-6) and the scoring guides, rated the same five essays independently, and then compared and discussed their ratings. They discussed how they interpreted the rating scales and which criteria they used in assigning their scores. Their discussion continued until they reached agreement in terms of both categories (Content & Organization and Language Use) of the writing. In fact, their ratings to those five essays were quite similar with exactly the same scores or a difference of one band apart. After the norming session, they went on with the rating independently at their own pace. They did the rating in the office for two hours, and then they switched the essays and took home the rest of the essays to finish.

The essays and the ratings were collected within a week. The two English raters reported that with band scores of 1-6 they found it hard to differentiate many of the essays. They found that many of the scores clustered around Band 3 (in effect, SPSS descriptive statistics shows that one of the raters marked Band 3 to 82 essays, out of a total of 136 essays). In addition, they hesitated sometimes if a score of .5 should be assigned. They suggested that it might be better to use a more detailed and differentiating rating scale, such as a scale of 15 points (with 1 as the lowest score and 15 as the highest). They said that they were more familiar with such a rating scale because it was adopted by large-scale standardized English tests in China such as College English Tests (CET-4 and CET-6) and Tsinghua English Proficiency Test (TEPT-I).⁴

The band rating scales were converted to 1-15 points (Appendix F lists both the band scale and the point scale). Two English teachers other than the first two raters were recruited in the same department to rate the English essays with the new point scale. New raters were sought in order to avoid repetition effect and bias from a previous reading of the essays from the “old”

⁴ CET-4 and CET-6 are nation-wide English tests for non-English major college students in China. More information can be obtained from their official website <http://www.cet.edu.cn/>. Tsinghua English Proficiency Test (TEPT-I) is for matriculated non-English major undergraduates at Tsinghua University as an exit exam for College English course. More information can be obtained from <http://www.tsinghua.edu.cn/docsn/wyx/chinese/ejiaoxue.htm> (general introduction in English) and <http://www.tsinghua.edu.cn/docsn/wyx/chinese/TEPT/tept1.htm> (detailed information in Chinese).

raters. As the first two raters, these two raters also had a lot of experience rating essays with such a point scale for CET and TEPT-I. They rated the essays together and compared and discussed their scores if the discrepancy was larger than three points. It took them over four hours to rate all the 136 copies of English essays.

The two Chinese essay raters were initially provided with the 1-6 band rating scale and scoring guides in Chinese. In the process of rating, the two Chinese raters also found the same problem of putting essays into bands and differentiating many of the essays. They were more used to rate on a scale of 1-100 points. Therefore, they converted the scale (see Appendix F for both the band scale and point scale for Chinese essays) and rated with the new scale of 1-100, with which they were more familiar and more experienced. The two Chinese raters also followed a similar norming session of rating five essays with discussion. During the rating process, they would compare their scores with each other. If their scores on the same essay differed by more than ten points, they would re-read the essay and try to give another score that would narrow the discrepancy. After rating all the essays with the new scale, they copied the scores on to the score sheet. It took the two Chinese raters two sessions with a total of approximately six hours to finish all the rating and conversion together.

The average of the scores assigned by two raters in each category was computed as the final score. Therefore, each essay has a Content score and a Language score and the maximum score for English essays was 15 points while that for Chinese essays was 100 points.

Use of Writing Strategies

Use of writing strategies in L1 and L2 were calculated separately based on the participants' responses to the questionnaire items. All the participants gave responses to the first 25 items regarding their familiarity of the genre and the topic as well as the strategies they used

before and during the writing of the essay. However, due to the fact that a number of participants answered “No” to the yes-no question preceding Section Four (about difficulties during writing) and Section Five (about revising) and therefore skipped the corresponding sections, there were a lot of missing data for items in those two sections. For statistical analysis, the items were grouped according to the conceptual sections. Specifically, scores were obtained by adding up the items in the relevant section. So there were scores for Genre/topic Knowledge (items 1-9), Planning Strategies (items 10-16), and Composing Strategies (items 17-25) for Chinese and English essay writing respectively. The other sections (Difficulty, Revision, and Use of L1) were analyzed separately due to missing data.

Working Memory Capacity

Total words scoring method was used in calculating the participant’s working memory capacity score, that is, the total number of the target words correctly recalled by a participant on each working memory task was recorded as his/her working memory capacity score for L1 and L2 respectively.

Traditionally, the span score is calculated as the maximum number of words that a participant recalls within a set.⁵ However, this way of scoring one’s reading span, a “quasi-absolute span score” in Conway et al.’s (2005) nomenclature, has serious drawbacks in terms of its reliability and distribution due to data loss/discarding, therefore it is not recommended (Conway et al., 2005; Friedman & Miyake, 2005). After comparing four scoring methods for the reading span test, Friedman and Miyake (2005) highly recommended the total words scoring method, because it yielded normal distributions and good reliability. Total words scoring method

⁵ According to Daneman and Carpenter (1980), the reading span test contained three sets of items, each with two to six sentences. The participants were presented with sets of increasingly more items until they failed all the three sets at a particular level. The level at which a participant recalled the target words correctly two out of three sets was taken as a measure of his/her reading span.

was also used by Van den Noort, Bosch, Haverkort and Hugdahl (2008) in their validation of a computerized reading span test they developed in four different languages.

For this project, the total number of correctly recalled target words was recorded as the score for a participant's working memory span. The order of the target words within a set was ignored. That is, the target words could be recalled in any order within the set they were shown. However, any target word that was recalled in a different set other than the set in which it appeared was counted as wrong (sometimes participants would recall and write down a target word that appeared in an earlier set). Any change in a letter of an English word or a stroke of a Chinese word was also counted as wrong. The total score for working memory span was 60 for both Chinese and English.

Vocabulary Scores

Two vocabulary scores were recorded according to the number of correct items the participants obtained in the vocabulary size tests—one for the receptive vocabulary test and the other for the controlled-production vocabulary test.⁶ The receptive vocabulary score was totally objective, because all the items are multiple choice questions. The productive vocabulary score was mostly objective, but for some of the answers, .5 scores were given if the word was correct but not the form (e.g., tense for verbs or plural for nouns). The total score for receptive vocabulary test was 130 and that for productive test was 90.

Grammar Score

For this computerized grammaticality judgment task, both the accuracy rate and the reaction time were recorded with the program DMDX. For the purpose of this study, a

⁶ It has to be noted that at the beginning of the receptive vocabulary test it says “wrong guesses will be taken away from your correct answers”. However, in the actual scoring, only corrected answers were counted without taking away wrong answers, on account of the following reasons: the participants were not alerted to this criterion and many of them did not realize that wrong answers would be penalized. In addition, some of them would have negative scores if such a penalty were implemented.

participant's grammar knowledge is reflected from his/her accuracy rate on the performance of this task. To be more exact, the number of accurate items was taken as a participant's grammar score. An accurate item referred to a participant's correct judgment of a sentence. That is, if a sentence is grammatical and was judged as grammatical, or if a sentence is ungrammatical and was judged as ungrammatical, it would be counted as an accurate item. If a participant's judgment was not congruent with the grammaticality of a sentence, the item would be counted as wrong. Those items recorded as "No response" (if the participant did not respond within the time limit) were counted as wrong. The total score for grammar test was 68.

Statistical Analysis

Variables

Six latent variables each represented by two or more indicators (observed variables) were involved in this study. Latent variables are variables that are not directly observed but are rather inferred from other observed variables. A total of 13 observed variables were used as indicators of the six latent variables. The following section presents the latent variables and their corresponding observations in detail.

L2 Writing Ability (EW)

L2 writing ability was measured by an English essay writing task. The latent variable L2 writing ability was represented by two indicators, observed as the analytical essay scores (i.e., Content & Organization and Language Use). Two experienced raters rated the English essays independently on an analytic scale of 1-15 points. The averages of the two scores for each category were used as the observed measures. The observed variables are English Content & Organization (EC) and English Language Use (EL). The measure for EW was obtained by taking the average of EC and EL.

L1 Writing Ability (CW)

L1 writing ability was measured by a Chinese essay writing task, similar to that for measuring L2 writing ability. The Chinese essays were rated by two experienced raters with a similar rubric on a scale of 1-100 points. The latent variable CW was thus represented by two observed variables: Chinese Content & Organization (CC) and Chinese Language Use (CL) obtained by the same procedure as for EC and EL. The measure for CW was the average of CC and CL.

L2 Language Knowledge (EK)

L2 language knowledge was represented by three variables measured by three instruments: a receptive vocabulary test (shortened as VOC1), a controlled-production vocabulary test (shortened as VOC2), and a timed grammaticality judgment task (shortened as GR). The measure for EK was obtained by adding up the scores for the three measures VOC1, VOC2 and GR.

Genre Knowledge (GK)

Genre Knowledge for Chinese (GKC) and Genre Knowledge for English (GKE) were obtained from the questionnaires following each of the timed essay writing tasks. Scores for the first nine items of each questionnaire were added up as measures for GKC and GKE respectively.

Strategy Use (SU)

The use of writing strategies for Chinese and English essay writing was also measured with the post-writing questionnaires separately. The measures for Chinese Strategy Use (CSU) and English Strategy Use (ESU) scores were obtained by adding up the sixteen corresponding items for planning and composing strategies for Chinese and English essay writing respectively.

Working Memory Capacity (WM)

Working memory capacity was measured by two separate tasks—operation span task in L1 (Chinese) and L2 (English)—with 60 items in each task, and therefore represented by two observed variables: working memory span in Chinese (WMC) and that in English (WME). The measure for WM was the average of WMC and WME.

As a summary, Table 3.2 presents the above mentioned variables and their corresponding observations.

Table 3.2 *Variables and Corresponding Observations*

Variables	Observations
L2 Writing Ability (EW)	English Content (EC)
	English Language (EL)
L1 Writing Ability (CW)	Chinese Content (CC)
	Chinese Language (CL)
L2 Language Knowledge (EK)	Receptive Vocabulary Knowledge (VOC1)
	Productive Vocabulary Knowledge (VOC2)
	Grammar Knowledge (GR)
Genre Knowledge (GK)	Genre Knowledge in Chinese (GKC)
	Genre Knowledge in English (GKE)
Strategy Use (SU)	Strategy Use in Chinese essay writing (CSU)
	Strategy Use in English essay writing (ESU)
Working Memory Capacity (WM)	Working Memory Capacity in Chinese (WMC)
	Working Memory Capacity in English (WME)

All the statistical analyses were performed using SPSS (version 17.0). Specifically, descriptive statistics, paired-samples *t*-test, analysis of variance (ANOVA), correlation, factor analysis, and multiple regression analysis were employed to explore the data.

CHAPTER 4

RESULTS

This chapter presents the results of the data analyses. The data analyses consisted of four major steps. First, demographic information of the participants was summarized. Second, descriptive statistical analyses of each observed variables were performed. Third, reliability checks and preliminary item analysis were carried out to ensure that the instruments were adequately reliable. Fourth, exploratory analyses were carried out with SPSS in an attempt to answer the research questions outlined in Chapter 1. More specifically, correlation, analysis of variance (ANOVA), paired-samples *t*-test, factor analysis and multiple regression analyses were performed to investigate the relationships between and among the variables.

Preliminary Data Analysis

Demographic Information of the Participants

In this section, detailed information about the participants' characteristics is presented. A total of 136 Chinese college students participated in this study. The participants came from four universities located in Beijing, China, with the vast majority (97%) from Tsinghua University, the host university for my data collection, and the others from three other universities.⁷ Table 4.1 presents the characteristics of the participants. The pool of the participants consisted of 104 undergraduates and 32 graduates from different majors/ departments. As can be seen, about 40% of the participants were engineering students and an additional 16.9% were studying sciences or computer science, which is representative of the student structure of the university. Among this sample group, 78 were male and 58 were female. Their age varied from 16 to 30 years old, with

⁷ As was explained in Chapter III, the data collection was conducted at Tsinghua University, where the majority of the participants were matriculated students. Those who came from other universities were recruited through the university's bulletin board system (BBS). They, with their access to the university's BBS, saw the advertisement for participants and contacted me via email.

a mean of 20 years and both a median and a mode of 19 years. A majority of the participants, 89.7% to be more exact, have been studying English for 6-12 years, and 70.6% of them started learning English at an age of 10-13 years old.

Table 4.1 *Characteristics of the Participants*

	Features	Number	Percent
Age	Mean: 20.18 Range: 16-30		
Age starting English	Mean: 10.82 Range: 3-15		
Years studying English	Mean: 9.40 Range: 6-23		
Gender	Male	78	57.4
	Female	58	42.6
Status	Undergraduate	104	76.5
	First year	75	
	Second year	16	
	Third year	2	
	Fourth year	11	
	Graduate	32	23.5
	First year	11	
	Second year	8	
	Third year	10	
	Fourth year	3	
Field of study	Art design	12	8.8
	Computer science	17	12.5
	Economics	4	2.9
	Engineering	55	40.4
	Humanities	28	20.6
	Law	6	4.4
	Psychology	1	.7
	Sciences	6	4.4
	Social sciences	7	5.1

According to the self-evaluation of their overall English proficiency, English writing ability and Chinese writing ability (see Table 4.2), over half of the participants rated their English proficiency level and their English writing ability as “Fair” (58.1% for the former and 62.5% for the latter, respectively), whereas over half of them rated their Chinese writing ability as either

“Good” (42.6%) or “Excellent” (11.8%).

Table 4.2 *Self-evaluation of English Proficiency Overall, English Writing Ability, and Chinese Writing Ability*

Self Evaluation	English Overall		English Writing		Chinese Writing	
	Frequency	Percent	Frequency	Percent	Frequency	Percent
Poor	25	18.4	25	18.4	11	8.1
Fair	79	58.1	85	62.5	51	37.5
Good	30	22.1	23	16.9	58	42.6
Excellent	2	1.5	3	2.2	16	11.8

There is an even more drastic difference in the participants’ self-report of hours spent on English writing per week versus those spent on Chinese writing (see Table 4.3). Fifty one point five percent (51.5%) of the participants reported that they spent less than an hour within a week writing English (including essays, letters, email, messages, journals/diary, etc.), and only 11.8% reported spending over three hours on English writing. In sharp contrast, those who spent less than an hour writing Chinese accounted for only 8.1% whereas those who spent more than three hours took up 55.1%.

Table 4.3 *Self-report of Hours per Week Spent on English and Chinese Writing*

Time spent per week	Writing English		Writing Chinese	
	Frequency	Percent	Frequency	Percent
< 1 hour	70	51.5	11	8.1
1-3 hours	50	36.8	50	36.8
3-5 hours	10	7.4	35	25.7
> 5 hours	6	4.4	40	29.4

Overall, the description of the background information of the participants suggests that the sample group appears to have included a wide range of college students at the host university for data collection.

Descriptive Statistics and Reliability of the Variables

Table 4.4 presents the descriptive statistics of the thirteen observed variables: English

Content (EC), English Language (EL), Chinese Content (CC), Chinese Language (CL), Working Memory Chinese (WMC), Working Memory English (WME), Grammar Knowledge (GR), Receptive Vocabulary Knowledge (VOC1), Productive Vocabulary Knowledge (VOC2), Genre Knowledge for Chinese essay writing (GKC) and for English essay writing (GKE), Strategy Use for Chinese essay writing (CSU) and for English essay writing (ESU). The means and standard deviations vary considerably among these variables due to the fact that different measures have different total scores.

Table 4.4 *Descriptive Statistics of the Observed Variables*

Measure	Minimum	Maximum	Maximum Possible Score	Mean	Std. Deviation
English Content (EC)	5.0	14.5	15	9.57	1.96
English Language (EL)	5.5	14.5	15	9.03	1.76
Chinese Content (CC)	45.0	89.0	100	69.48	9.52
Chinese Language (CL)	52.5	87.0	100	69.11	8.10
Working Memory Chinese (WMC)	34.0	60.0	60	55.60	4.46
Working Memory English (WME)	36.0	60.0	60	51.47	5.32
Grammar (GR) all	35.0	62.0	68	48.41	6.04
Vocabulary Receptive (VOC1) all	15.0	84.0	130	35.65	12.42
Vocabulary Productive (VOC2) all	11.0	67.5	90	30.82	10.40
Genre Knowledge Chinese (GKC)	11.0	45.0	45	29.85	6.36
Genre Knowledge English (GKE)	13.0	45.0	45	30.87	6.08
Chinese Strategy Use (CSU)	24.0	77.0	80	57.06	8.58
English Strategy Use (ESU)	24.0	77.0	80	55.14	8.21

Table 4.5 presents the reliability measures and the means and variances of the percentage accuracy scores for each of the measures for working memory spans (WMC and WME), L2 language knowledge (GR, VOC1, VOC2), and genre knowledge (GKC and GKE), and use of writing strategies (CSU and ESU). The item mean statistics shows that for instruments that used one point for each item, the participants obtained the highest mean score on Working Memory Chinese ($M = .923$) and the lowest mean score on Vocabulary Receptive ($M = .257$). The item

variance statistics, which reflects the dispersion of the distribution of scores, ranges from .067 for Working Memory Chinese and .182 for Grammar. For questionnaire items that used five as the highest scale for each item, the item means and variances are very similar (around 3.4 for item mean and around 1.3 for item variance). Cronbach's alphas were calculated as estimates of internal consistency for each of the measures. The reliability estimates vary from .731 for English Strategy Use to .916 for Vocabulary Receptive, demonstrating acceptably good reliability, except for Grammar (.671).

Table 4.5 *Reliability and Item Statistics*

Measure	Number of Items	Item Mean	Item Variance	Reliability
Working Memory Chinese	60 ¹	.923	.067	$\alpha = .824$
Working Memory English	60 ²	.855	.110	$\alpha = .783$
Grammar all	68	.712	.182	$\alpha = .671$
Vocabulary Receptive all	130 ¹	.257	.111	$\alpha = .916$
Vocabulary Productive all	90 ¹	.354	.121	$\alpha = .913$
Genre Knowledge Chinese	9	3.317	1.358	$\alpha = .785$
Genre Knowledge English	9	3.430	1.284	$\alpha = .773$
Chinese Strategy Use	16	3.566	1.260	$\alpha = .795$
English Strategy Use	16	3.446	1.325	$\alpha = .731$

1. Three items have zero variance and are removed from the scale when SPSS calculates the reliability statistics.
2. One item has zero variance and is removed from the scale when SPSS calculates the reliability statistics.

Item analyses were carried out for each of the measures in order to improve the reliability estimates by detecting and deleting items responsible for low reliability. A closer examination of the item-total statistics (e.g., item-total correlation and Cronbach's alpha if item deleted) was made and items with a negative item-total correlation were eliminated or modified by reversing the scale. For the measures for working memory spans (WMC and WME), almost all the items were performing normally; therefore, no item was deleted for reanalysis. For the measures for use of writing strategies (CSU and ESU), the scale for one item (the same item for both the Chinese and the English questionnaire) was reversed before the analysis because that item was asking the same question as the preceding item in an opposite way. With this treatment, the items

were all performing normally and the reliability measures were acceptably high. For the measurement of Grammar, 10 out of the 68 items were found to have a negative value for corrected item-total correlation. When these 10 items were excluded in the reliability analysis, the Cronbach's alpha was improved to .711, which is acceptably moderate.⁸

For the two vocabulary measures (receptive and productive vocabulary), even though the alpha values were sufficiently high, the item means were very low. Closer examinations of the answer sheets revealed that a large number of participants did not answer the second half of the receptive test and the last batch of the productive test. This phenomenon deserves some explanation here, so as to justify the adjustment of scores for the two vocabulary measures. The low scores in vocabulary tests could be explained by the structure of the test which includes increasingly difficult items up to the vocabulary level of 13,000 for the receptive test and 10,000 for the productive test. In fact, in the receptive test, about 19% of the participants were not able to complete past Item 70, and for those who did get beyond Item 70 to some of the remaining 60 items, 51% out of the total participants got only 1-5 answers correct. In the productive test, about 66% students were not able to get past the University Word List level. Therefore, the remaining items were dropped and the participants' vocabulary scores were recalculated, using the first 70 items in the receptive test (which tested the receptive vocabulary size at 7000 level) and the first 72 items in the productive test (which tested the productive vocabulary size at the University Word List level).

⁸ Compared with R. Ellis (2005, $\alpha = .81$), the reliability for the timed GJT in this study ($\alpha = .671$) was rather low. Even though the reliability was improved to .711 after removing negatively performing items, it was still much lower than that in Ellis' study. One possible reason might be participants' familiarity, or rather, lack of familiarity, with this type of task. All the participants commented that this was the first time they ever did such a grammaticality judgment task on a computer, and some of them reported that they were so nervous, especially at the beginning of the task, that they would hit the wrong key by mistake. This phenomenon might have brought about more random error into the outcome of this task, thus influencing negatively the reliability of this measure.

Table 4.6 presents the reliability estimates and item means and item variances of each of the three measures after adjustment. The reliability estimates ranged from .711 for GR and .902 for VOC2, demonstrating moderate to high reliability for each of the adjusted measures. The statistics for item means and item variances were also improved. Table 4.7 presents the descriptive statistics of these three measures after adjustment.

Table 4.6 *Reliability and Item Statistics of Measures after Adjustment*

Measure	Number of Items	Item Mean	Item Variance	Reliability
Grammar	58	.700	.189	$\alpha = .711$
Vocabulary Receptive	70 ¹	.412	.145	$\alpha = .852$
Vocabulary Productive	72	.419	.139	$\alpha = .902$

1. Three items have zero variance and are removed from the scale when SPSS calculates the reliability statistics.

Table 4.7 *Descriptive Statistics of Measures after Adjustment*

Measure	Minimum	Maximum	Maximum Possible Score	Mean	Std. Deviation
Grammar	27.0	55.0	58	40.59	6.03
Vocabulary Receptive	15.0	57.0	70	30.63	7.78
Vocabulary Productive	11.0	59.5	72	30.19	9.50

Inter-rater Reliability

Due to the subjective nature of essay rating, a source of error typically lies in the inconsistency of ratings of essays. Therefore, inter-rater reliability was calculated to measure the consistency of the two raters for the Chinese and English essays respectively. Pearson correlation coefficients were computed and reported in Table 4.8 for Chinese essay raters and Table 4.9 for English essay raters. As can be seen from the tables, the inter-rater reliability estimates among the two raters for the Chinese essays and the two for the English essays were found to be sufficiently high, ranging from .845 for English Content to .918 for Chinese Language. No exact or adjacent agreement percentage statistics was reported here because the

scales for essay rating were not band scores, but point scores (1-100 for Chinese essays and 1-15 for English essays).

Table 4.8 *Inter-rater Reliabilities for Chinese Essay Ratings*

Variable	Chinese Language rater 1	Chinese Content rater 2	Chinese Language rater 2
Chinese Content rater 1	.853**	.892**	.813**
Chinese Language rater 1		.791**	.918**
Chinese Content rater 2			.857**

Table 4.9 *Inter-rater Reliabilities for English Essay Ratings*

Variable	English Language rater 1	English Content rater 2	English Language rater 2
English Content rater 1	.858**	.845**	.798**
English Language rater 1		.782**	.853**
English Content rater 2			.851**

Exploratory Analyses of the Relationships among the Variables

In order to answer the research questions set forth for this study, a series of statistical analyses were carried out to explore the relationships between and among the variables investigated. As a first step, a correlation matrix for all the thirteen observed variables was created (see Appendix G). In the following sections, relationships among these variables were investigated in an attempt to answer the specific research questions.

Research Question 1: Relationship between Writing Ability and Working Memory Capacity

Table 4.10 shows the correlations between variables for L1, L2 writing ability (Content and Language for Chinese and English writing respectively) and working memory capacity in Chinese and English (WMC and WME). From the correlation coefficients we can see that the two measures for working memory capacity in the two languages correlate with each other

moderately ($r = .594$). However, no correlation was found between Working Memory in Chinese (WMC) and Chinese writing (CC and CL) or between Working Memory in English (WME) and English writing (EC and EL). A closer examination of the scatter plots revealed that most of the scores for WMC clustered around the higher end of the maximum possible scores. This distribution indicates a possible ceiling effect of the measurement for WMC. These results will be discussed in detail later in Chapter V.

Table 4.10 *Correlations between Working Memory Capacity and Writing Scores in L1 and L2*

Variable		Working Memory Capacity	
		English	Chinese
English Writing	Content	.161	.040
	Language	.114	.034
Chinese Writing	Content	.208*	.034
	Language	.262**	.017
Working Memory Capacity	English	-	.594**
	Chinese		-

In terms of the relationship between the working memory capacity in L1 and L2, a paired-samples t -test was conducted on the number of words recalled in Chinese and English. Results show that participants recalled significantly more words in Chinese than in English ($t = 10.794$, $df = 135$, $p = .000$).⁹ Therefore, it can be concluded that the functional working memory capacity of the participants was larger in their L1 compared to their L2.

Research Question 2: Relationship between L1, L2 Writing Ability and L2 Language Knowledge

Table 4.11 shows the correlations between the two variables for L1 writing ability (CC and CL) and those for L2 writing ability (EC and EL). The Content and Language scores for the same language are significantly highly correlated, with $r = .870$ for CC and CL and $.889$ for EC

⁹ Means and standard deviations of working memory capacity in Chinese and English were reported earlier in Table 4.4. For easier reference, the descriptive statistics was repeated here: for WMC, $M = 55.60$, $SD = 4.46$; for WME, $M = 51.46$, $SD = 5.32$.

and EL. However, across the languages, the correlation coefficients are not significant for EL, either with CC or CL. And for EC, although the correlation coefficients with CC and CL are statistically significant, the magnitudes are rather low ($r = .194$ and $.277$ respectively).

Table 4.11 *Correlations between Content and Language Scores in L1 and L2*

Variable	Chinese Content	Chinese Language	English Content	English Language
Chinese Content	-	.870 ^{**}	.194 [*]	.050
Chinese Language		-	.277 ^{**}	.124
English Content			-	.889 ^{**}
English Language				-

In an attempt to further explore the relationship between L1 and L2 writing ability, the two latent variables for L1 and L2 writing ability were examined. Chinese Writing Score (representing L1 writing ability) and English Writing Score (representing L2 writing ability) were generated by computing the average of the Content and Language scores for each language. The correlation between these two variables was found to be very weak yet significant at $\alpha = .05$ level ($r = .172, p = .045$). A closer examination of the scatter plots of Chinese Writing Score and English Writing Score revealed a pattern of higher scores in both languages tending to cluster together, which suggests different patterns for different levels. In order to investigate the differences among the levels of their Chinese writing and its relationship with English writing, the participants' Chinese and English Writing Scores were further grouped into levels according to band scales (see Appendix F for the conversion of rating scales). The levels for Chinese writing ranged from 2 (poor) to 5 (very good) and the levels for English writing ranged from 3 (fair) to 6 (excellent). Table 4.12 shows the crosstabulation of the number of band scores that fall into each level. It can be seen from the table that none of those who were poor or fair in the Chinese essay (Band 2 and 3) got an excellent score in the English essay (Band 6), whereas none

of those who got a very good score in the Chinese essay (Band 5) were poor in the English essay (Band 3). In addition, those who got an excellent score in the English essay (Band 6) all got good or very good scores in the Chinese essay (Band 4 and 5). In other words, participants who are good writers in Chinese tend to be better writers in English.

Table 4.12 *Crosstabulation of Chinese Writing Band and English Writing Band*

		English Writing Band				Total
		3	4	5	6	
Chinese Writing Band	2	4	15	2	0	21
	3	5	29	8	0	42
	4	10	29	13	1	53
	5	0	10	8	2	20
Total		19	83	31	3	136

Table 4.13 presents the descriptive statistics of English Writing Score across the Chinese Writing Band. Analysis of variance (ANOVA) with English Writing Score as the dependent variable and Chinese Writing Band as the independent variable was performed to see if there are significant differences among the groups. Levene's test shows that the error variance of the dependent variable is equal across groups ($F_{(3, 132)} = 1.099, p = .352$). ANOVA results suggest that there is a significant difference among the level groups ($F_{(3, 132)} = 3.183, p = .026$). Post hoc tests with Scheffe revealed that the English Writing Scores of Group 2 (lowest in Chinese Writing) and Group 5 (highest in Chinese Writing) are significantly different, which indicates that those who are poor Chinese writers produced poor English writing, whereas those who are good Chinese writers produced significantly better English writing.

Table 4.13 *Descriptive Statistics of English Writing Scores across Chinese Writing Bands*

Dependent Variable: English Writing Score			
Chinese Writing Band	Mean	Std. Deviation	N
2	8.67	1.59	21
3	9.19	1.57	42
4	9.27	1.99	53
5	10.31	1.70	20
Total	9.30	1.81	136

In summary, statistical analysis revealed low correlations between English Content and Chinese writing (both with Content and Language) but no correlations between English Language and Chinese writing. Closer examinations of level groups showed that those who were poor in the Chinese essay writing did not perform well in the English essay writing, whereas those who performed well in English seemed to be good writers in Chinese as well.

The relationships among L1 writing ability, L2 writing ability, and L2 language knowledge were investigated from the perspective of linguistic threshold hypothesis (Cummins, 1979), which suggests that learners' L1 skills might be able to transfer into L2 only when their L2 proficiency reaches a certain threshold level. In regard to this hypothesis, the relationship between L1 and L2 writing ability was reexamined across different levels of L2 language knowledge.

For this analysis, the participants were divided into three subgroups with similar sample size, according to the level of L2 language knowledge as represented by the latent variable English Knowledge (EK, a combination of grammar and vocabulary scores). Table 4.14 presents the means and standard deviations of EK, Chinese Writing Score (CW) and English Writing Score (EW) across the EK level groups, as well as the Pearson correlation coefficients of CW and EW for each EK level group. It is interesting to note that the correlation between L1 and L2

writing scores tended to increase as the level of L2 language knowledge increased, though none of the correlations reached the significance level.

Table 4.14 *Relationship among EK, CW and EW across EK levels*

EK level	N	EK		CW		EW		$r_{(CW, EW)}$
		M	SD	M	SD	M	SD	
Low	43	80.61	8.20	67.52	8.14	8.62	1.54	.064
Medium	50	101.10	4.60	69.44	8.51	8.98	1.53	.138
High	43	122.58	11.06	70.90	8.75	10.35	1.91	.162

EK = English Knowledge, CW = Chinese Writing Score, EW = English Writing Score

One-way ANOVA was performed to compare the writing scores across the English level groups in the two languages respectively. Levene's test shows that both the Chinese and the English Writing Scores have homogeneity of variances across the three English level groups ($F_{(2, 133)} = 2.305, p = .104$ for English, $F_{(2, 133)} = .072, p = .931$ for Chinese). ANOVA results suggest that there is a significant difference among the level groups for English Writing Scores ($F_{(2, 133)} = 13.195, p = .000$) but no significant difference for Chinese Writing Scores ($F_{(2, 133)} = 1.726, p = .182$). Post hoc tests with Scheffe revealed that the English Writing Scores of high English level group differ significantly from those of low and medium English level groups, which indicates that participants with more English knowledge (higher English proficiency) performed significantly better in English essay writing but not necessarily in Chinese essay writing.

The relationship between L2 writing performance and L2 language knowledge was further explored by examining the correlation coefficients between the observed variables for these two facets. Table 4.15 shows the correlation coefficients between these variables for L2 writing ability (EC and EL) and L2 language knowledge (GR, VOC1, and VOC2). All the correlations are found to be significant at $\alpha = .01$ level, except for the pair of VOC1 & GR. Receptive and Productive Vocabulary Knowledge shows a high correlation of .698, whereas the magnitude of the other pairs of correlations ranges from low to medium, suggesting not so strong

correlations among these variables. Among the three observed variables for L2 language knowledge, Productive Vocabulary Knowledge (VOC2) has the highest correlations with EC ($r = .419$) and EL ($r = .446$). Linear regression analyses with EC and EL as the dependent variable respectively and the three predictors for L2 language knowledge as the independent variables show that 19.1% of the variance of EC and 22.1% of the variance of EL can be explained by the three independent variables ($R^2 = .191$ for EC and $.221$ for EL). Among the three independent variables, VOC2 has the largest explanatory power for both EC and EL, whereas the impacts of GR and VOC1 are insignificant and negligible (see Table 4.16).

Table 4.15 *Correlations between Subscales for L2 writing and L2 language knowledge*

Variable	Grammar (GR)	Vocabulary Receptive (VOC1)	Vocabulary Productive (VOC2)
English Content	.275 ^{**}	.303 ^{**}	.419 ^{**}
English Language	.315 ^{**}	.291 ^{**}	.446 ^{**}
Grammar		.123	.400 ^{**}
Vocabulary Receptive			.698 ^{**}

Table 4.16 *Regression Summary for GR, VOC1 and VOC2 Predicting EC and EL*

Predictors	English Content (EC)		English Language (EL)	
	R ²	Standardized coefficients	R ²	Standardized coefficients
Vocabulary Productive (VOC2)	.175	.320	.199	.373
Grammar (GR)	.014	.139	.022	.165
Vocabulary Receptive (VOC1)	.002	.062	.000	.011

In summary, results show that English Content is slightly correlated with Chinese content and Chinese Language whereas English Language is not with either of the two. Both English Content and Language correlate with measures for English language knowledge, namely, grammar, receptive and productive vocabulary knowledge, and productive vocabulary knowledge seems to be the best predictor for English writing scores among the three measures.

Research Questions 3: Relationship between L1, L2 Writing and the Use of Strategies

In this section, the similarities and differences in the use of strategies in L1 and L2 essay writing will be examined, as well as the relationships between strategy use and writing performance in L1 and L2. First, correlation coefficients between the writing scores and the scores for Genre Knowledge and Strategy Use in both languages are presented. Then, results from the examination of the questionnaire items are presented in terms of a) genre knowledge and familiarity to the task, b) planning and composing strategies used in L1 and L2 essay writing, c) difficulties faced in L1 and L2 writing process, d) revision strategies used in L1 and L2 writing, and e) use of L1 in L2 writing.

As explained earlier in Chapter 3, the scores for Genre Knowledge and Strategy Use for both L1 and L2 writing were obtained by adding up the ratings of the corresponding items in the corresponding questionnaire. Table 4.17 presents the correlation coefficients between the observed variables for L1 writing (CC and CL), L2 writing (EC and EL), Genre Knowledge for L1 (GKC) and L2 (GKE), and Strategy Use for L1 (CSU) and L2 (ESU).

Table 4.17 *Correlations between Subscales for Strategy Use and Writing Performance*

Variables		Genre Knowledge		Strategy Use	
		English	Chinese	English	Chinese
English Writing	Content	.357**	.099	.318**	.111
	Language	.357**	.117	.273**	.066
Chinese Writing	Content	.113	.293**	.031	.162*
	Language	.149	.232**	.082	.197*
Genre Knowledge	English	-	.296**	.474**	.177*
	Chinese		-	.182*	.522**
Strategy Use	English			-	.375**
	Chinese				-

The results show many pairs of low but significant correlations. Specifically, Genre Knowledge and Strategy Use correlates with each other in both languages; Genre Knowledge

and Strategy Use in English correlate with Content and Language in English but not with those in Chinese, whereas Genre Knowledge and Strategy Use in Chinese correlate with Content and Language in Chinese but not with those in English. It is also interesting to note that the magnitude of the correlations among the variables for English writing is greater than the magnitude of those for Chinese writing.

Next, the questionnaire data were put into a closer examination in order to answer the secondary research question concerning the similarities and differences between writing in L1 and L2. Paired-samples *t*-tests were performed for all the matching items in the two questionnaires. The results showed significant differences for some of the items. In the following paragraphs, I will first report the results for the items about genre knowledge and strategy use in planning and composing, because these items elicited complete data from all the participants. Then I will present the results for the items about difficulties faced and revision strategies used in the writing process, which elicited data from some of the participants.

Table 4.18 presents the items that are significantly different in Chinese and English essay writing in terms of Genre Knowledge and Strategy Use. Participants reported that they had more practice writing this type of argumentative essays in English (#3) and they had practiced writing this topic or a similar topic in English (#9). While writing the English essay, they would use a ready template more often (#5); they would take some notes of the main arguments and/or keywords before they started (#14), and they would state the main points in the first paragraph (#17). In contrast, while writing the Chinese essay, they seemed to have more ideas for the topic (#8); they tended to do less planning but start writing without a written or mental outline (#16); they wrote more paragraphs and longer essays (#18); and they tried to write coherently (#24).

However, the effect sizes of the t -statistic for these items ranged from small to medium,¹⁰ which indicates that the differences, despite their statistical significance, were not very large.

Table 4.18 *Items Showing Significant Differences in Chinese versus English Genre Knowledge and Strategy Use*

	Items	t^1	Mean ²	Effect Size	Sig. (2-tailed)
Genre Knowledge	3. I have practiced writing this type of essay	-2.32	E > C	.28	.022
	5. I used a ready framework/ template in my writing	-2.22	E > C	.27	.028
	8. I had enough ideas to write about this topic.	3.20	C > E	.38	.002
	9. I have practiced writing this topic or a similar topic	-3.99	E > C	.48	.000
Planning Strategy	14. Before I started writing, I took some notes such as main arguments and keywords	-2.64	E > C	.32	.009
	16. I started writing the essay without a written or a mental outline	2.89	C > E	.35	.005
Composing Strategy	17. I stated my main points in the first paragraph	-2.12	E > C	.25	.036
	18. I wrote at least three paragraphs	2.11	C > E	.25	.037
	24. I tried to connect my ideas smoothly	2.20	C > E	.26	.030

1. $df = 135$

2. E > C indicates the mean of the item for English writing is larger than that for Chinese writing; C > E indicates the mean of the item for Chinese writing is larger than that for English writing.

In terms of the difficulties faced and the revision strategies used in the writing process, two Yes-No questions were included in both questionnaires to elicit whether the participants had difficulty with word choice during the writing process (#26: I sometimes had difficulties when I tried to express my ideas with appropriate words.) and whether they did any revision to their essays (#33: I spent some time reading back and revising what I had written down). If their answer to either question was “Yes,” they were directed to answer more questions regarding their strategies for dealing with the difficulties and for revision. Table 4.19 presents the frequencies and percentages of the answers to the two Yes-No questions (Difficulty for Item #26 and Revision for Item #33).

¹⁰ According to Stevens (2007), for t -test, an effect size of .20 is small, .50 medium, and .80 large.

Table 4.19 *Frequencies and Percentages of Yes-No Answers to Item #26 & #33*

	Chinese Difficulty		Chinese Revision		English Difficulty		English Revision	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	percent
Yes	49	36.0	72	52.9	133	97.8	97	71.3
No	87	64.0	64	47.1	3	2.2	39	28.7

Figure 4.1 shows the sharp contrast between the number of those who answered Yes and those who answered No to these two questions while writing in the two different languages. More than half participants reported having no difficulty writing the Chinese essay whereas a vast majority of the same participants reported having difficulty writing the English essay. In terms of revision, about half reported re-reading and revising while writing the Chinese essay, as opposed to about two thirds while writing the English essay.

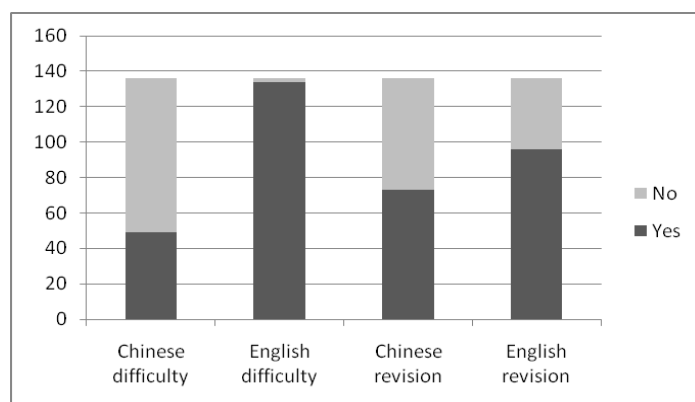


Figure 4.1 Contrast of Numbers of Yes/No Answers to Difficulty and Revision

Paired-samples *t*-tests were also performed for Difficulty items and Revision items using available data.¹¹ The results showed that none of the paired items for Difficulty was significantly different and only five out of the 16 items for Revision were significantly different. For strategies used to cope with difficulties encountered in the writing process, the participants reported similar strategies for both L1 and L2 such as pausing to find a solution (#27) rather than leave it until

¹¹ For Difficulty items, 48 cases had data in both Chinese and English questionnaires, and for Revision items, the number was 63.

finishing the whole essay (#28), using a similar English word (#29) or finding another way to express the same idea (#30) rather than give up the idea and write something else (#31).

Table 4.20 presents the statistics for the five statistically different items for revision strategies. It seems that during the revision stage in the English essay writing, participants did more online revision (#37), that is, revising while writing instead of after the completion of the essay, and they focused more on local aspects such as grammar, spelling, and punctuation (#41, 42, and 43), whereas while writing the Chinese essay, they focused more on global aspects such as content and meaning (#39). The medium to large effect size for the *t*-statistic indicates that the differences are considerable, especially for #41 and #43.

Table 4.20 *Items Showing Significant Differences in Chinese versus English Revision*

Items		<i>t</i> ¹	Mean ²	Effect Size	Sig. (2-tailed)
Revision Strategies	37. I revised as I wrote throughout the writing of the entire essay.	-3.31	E > C	.60	.002
	39. I focused mostly on content and meaning.	4.00	C > E	.72	.000
	41. I focused mostly on grammar, spelling, and punctuation.	-5.31	E > C	.96	.000
	42. I focused on both content and grammar.	-2.68	E > C	.48	.010
	43. I corrected spelling and/or grammatical mistakes.	-4.77	E > C	.86	.000

1. *df* = 62

2. E > C indicates the mean of the item for English writing is larger than that for Chinese writing; C > E indicates the mean of the item for Chinese writing is larger than that for English writing.

In order to investigate the use of L1 in the writing of L2 essays, five more items were included in the questionnaire for the English essay writing. Table 4.21 presents the descriptive statistics and frequencies of each answer to these five items. It seems obvious that the majority of the participants did not use Chinese word(s) in place of an English word when they had difficulty finding an appropriate word (#33). Rather, they would use other strategies for coping with difficulties such as those reported earlier. Items 51-54 indicate that most students used both

languages while writing the English essay (#52 & 53), only a few (N = 38) relied on Chinese for the initiation of ideas and used translation to produce the English essay, and even fewer students (N = 26) used English only while writing the English essay. It can be concluded that most participants used both languages in the process of writing the English essay.

Table 4.21 *Use of L1 in L2 Essay Writing*

Items	N	Mean	Std. Deviation	Frequency					
				0*	1	2	3	4	5
eq33 I used Chinese and later tried to find an appropriate English word.	133	1.59	.99	6	72	38	7	8	2
eq51 I formulated my ideas mostly in Chinese and translated them into English.	134	2.62	1.40	4	28	39	25	20	18
eq52 I generated ideas mainly in Chinese, but formulated sentences in English.	134	3.04	1.37	3	19	28	24	40	20
eq53 I generated ideas in both languages, but formulated sentences in English.	134	2.91	1.34	5	17	32	23	45	12
eq54 I thought mostly in English and wrote in English.	134	2.13	1.31	8	41	44	15	18	8

*0 = not applicable; 1 = strongly disagree; 2 = disagree; 3 = not sure; 4 = agree; 5 = strongly agree

It is also my interest to find out if the participants' writing performance in English was related to the revision strategies used while writing the English essay. Further analyses were carried out with the 16 items concerning the revision strategies for the English essay. First, reliability of the 16 items was checked by calculating the Cronbach's alpha. The value of the Cronbach's alpha for the original 16 items is .624. However, #36 and #38 are the opposite statements of #35 and #37 respectively. Therefore, the scales for these two items were reversed, and the Cronbach's alpha was improved to .781.

Next, a factor analysis was conducted for the 16 items using Principal Component Analysis with Varimax rotation. Five components were extracted (see Table 4.22) with some items loaded on more than one factor.

Table 4.22 *Rotated Component Matrix for Revision Items*

Rotated Component Matrix ^a					
	Component				
	1	2	3	4	5
eq36 reversed	.852				
eq38 reversed	.819				
eq35	.704				
eq37	.701				
eq46		.815			
eq47		.813			
eq45		.685			
eq39			.755		
eq48			.742		
eq42			.642		
eq49			.548		.467
eq40			.485		
eq43				.807	
eq41				.673	
eq44				.620	
eq50					.863

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 6 iterations.

According to the value in the Rotated Component Matrix and the focus of each item, these 16 items were divided into five categories:

Item number	Component Focus
35-38	Revise Time (while writing or after writing)
39, 40, 42, 48, 49	Global Focus
45-47	Global Change
41, 43, 44	Local Focus/ Change
50	Audience Awareness

Five new variables were created by combining the items in each of the components. Pearson correlation coefficients were calculated for these variables and the two for L2 writing (English Content and English Language). However, none of the correlations was found to be statistically

significant (see Table 4.23).

Table 4.23 *Correlations between EC / EL and Revision Strategies*

	English Revise Time	English Global Focus	English Global Change	English Local Change	English Audience Awareness
English Content	.078	.100	.112	.034	.024
English Language	.127	.174	.128	.027	.079

In summary, analysis of questionnaire data yielded a complicated picture of the relationship between strategy use and writing performance in L1 and L2. Both similarities and differences have been found in the strategy use for writing L1 and L2 essays, and writing performance in L2 seems to be correlated with some factors (such as genre knowledge and planning strategies) but not with others (such as revision strategies).

Research Question 4: Relative Importance of the Selected Variables to L2 Writing Performance

In this section, the relationship between L2 writing ability (as the dependent variable) and other factors (as the independent variables) would be examined using multiple regression analyses with an aim to learn more about which variables could be better predictors of L2 writing ability. First, observed variables were grouped into latent variables in order to avoid the problem of multicollinearity in the process of multiple regressions. Second, multiple regression procedures were performed to generate models for the relative importance and strength of associations of the predictors in relation to the dependent variable.

Independent Variables as Predictors

A big concern in multiple regression analysis is multicollinearity, which refers to the existence of strong correlations between independent variables. If two variables are significantly highly correlated, it would be extremely difficult to determine which of the variables account for the variance in the dependent variable. Therefore, the existence of multicollinearity may

influence the magnitude of parameter estimates and result in incorrect conclusions about relationships between independent and dependent variables in multivariate analysis (von Eye & Schuster, 1998). The correlation matrix in Appendix G shows that various degrees of correlations exist between the observed variables and that some pairs have correlations over .80 (for example, .889 between English Content and Language, .870 between Chinese Content and Language), which indicates a possible problem of multicollinearity according to Hatch and Lazaraton (1991). In addition, too many independent variables tend to inflate the variability of the predicted value and thus lower the precision of the prediction (von Eye & Schuster, 1998). Therefore, variables measuring the same latent variable were combined (English Content and Language were combined as English Writing Score, Chinese Content and Language as Chinese Writing Score, Grammar, Receptive Vocabulary and Productive Vocabulary as English Knowledge). Five variables were selected as the predictors for English writing ability—English Knowledge, Chinese Writing Score, Working Memory Span, Strategy Use and Genre Knowledge. Scores for each independent variable were computed by taking the average or the sum of the composing variables.¹² The descriptive statistics for each of the latent variables and their correlation coefficients were reported below (see Table 4.24 and 4.25). In subsequent analyses, these latent variables were used for multiple regression procedures.

¹² For English Writing Score, Chinese Writing Score, and Working Memory Span, the averages of EC and EL, CC and CL, and WMC and WME were computed respectively. For English Knowledge, the sum of GR, VOC1 and VOC2 was computed. For Genre Knowledge and Strategy Use, the scores obtained from the questionnaire items for English essay writing were used rather than taking either the average or the sum of the scores for English and Chinese Writing Strategies, because it would make no sense combining the two scores considering the differences between some of the items in the two measures as illustrated in a previous section.

Table 4.24 *Descriptive Statistics of the Latent Variables*

	N	Minimum	Maximum	Maximum Possible Score	Mean	Std. Deviation
English Writing Score	136	5.25	14.50	15	9.30	1.81
Chinese Writing Score	136	52.00	88.00	100	69.29	8.52
Working Memory Span	136	36.5	60.0	60	53.53	4.37
English Knowledge	136	59.5	153.5	200	101.41	18.64
Genre Knowledge	136	13.0	45.0	45	30.87	6.08
Strategy Use	136	24.0	77.0	80	55.14	8.21
Valid N (listwise)	136					

Table 4.25 *Pearson Correlation Coefficients of the Latent Variables*

	Chinese Writing Score	Working Memory Span	English Knowledge	Genre Knowledge	Strategy Use
English Writing Score	.172*	.106	.452**	.368**	.305**
Chinese Writing Score		.160	.169*	.134	.056
Working Memory Span			.177*	-.064	-.059
English Knowledge				.235**	.119
Genre Knowledge					.474**

Multiple regression analyses

A stepwise linear regression analysis was first performed with English Writing Score as the dependent variable and the five factors (Chinese Writing Score, English Knowledge, Genre Knowledge, Strategy Use and Working Memory Span) as the independent variables. In a stepwise regression, models are tested with one more variable at a time. SPSS first tests a model with the most correlated independent variable. Then it tests a model with this most correlated variable plus the variable with the highest partial correlation with the dependent variable controlling for the most correlated independent variable. Those variables that do not significantly increase R-squared are excluded from model building and models with these variables are not considered (Bryman & Cramer, 2009).

Stepwise regression results show that three variables English Knowledge (EK), Genre Knowledge (GKE) and Strategy Use (ESU) contribute significantly to the dependent variable English Writing Score, with EK being the most important predictor, explaining approximately 20.4% of the variance, and GKE adding 7.2%, and ESU adding another 2.1% (see Table 4.26). The other two variables Chinese Writing Score (CW) and Working Memory Span (WM) were excluded in the process of model building, suggesting that these two variables do not have a sufficient impact on English Writing Score.

Table 4.26 *Model Summary from Stepwise Multiple Regression*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.452 ^a	.204	.198	1.621	.204	34.357	1	134	.000
2	.526 ^b	.277	.266	1.551	.072	13.318	1	133	.000
3	.546 ^c	.298	.282	1.534	.021	3.977	1	132	.048

a. Predictors: (Constant), English Knowledge

b. Predictors: (Constant), English Knowledge, Genre Knowledge English

c. Predictors: (Constant), English Knowledge, Genre Knowledge English, English Strategy Use

d. Dependent Variable: English Writing Score

In order to estimate how much contribution each of the independent variables makes to the dependent variable, the multiple regression analysis was performed again with the same dependent and independent variables with the method “Backward” rather than “Stepwise.” This way of running multiple regression procedure makes sure that all the independent variables would be entered in model building and that R Square Change and Beta Weight¹³ would be calculated for each of the independent variables.

Table 4.27 shows the model summary of the backward multiple regression procedure.

The overall model was found to be significant ($F = 11.443$, $p = .000$), with $R^2 = .306$, explaining

¹³ Beta Weight, also known as “standardized regression coefficient”, tells us how many standard deviation units the dependent variable will change for a one-standard-deviation change in the independent variable. It standardizes the units of measurement involved in the independent variables and thus makes possible the comparison of the relative importance of the predictors involved.

30.6% of the variance in the dependent variable. Change statistics indicates that Working Memory Span explains only 0.3% of the variance and Chinese Writing Score explains 0.5%. The impacts of these two variables were so small that they were almost negligible and were therefore excluded in the stepwise regression.

Table 4.27 *Model Summary from Backward Multiple Regression*

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.553 ^a	.306	.279	1.537	.306	11.443	5	130	.000
2	.550 ^b	.303	.282	1.534	-.003	.499	1	130	.481
3	.546 ^c	.298	.282	1.534	-.005	.992	1	131	.321

a. Predictors: (Constant), English Strategy Use, Chinese Writing Score, Working Memory, English Knowledge, Genre Knowledge English

b. Predictors: (Constant), English Strategy Use, Chinese Writing Score, English Knowledge, Genre Knowledge English

c. Predictors: (Constant), English Strategy Use, English Knowledge, Genre Knowledge English

d. Dependent Variable: English Writing Score

Table 4.28 *Coefficients from Backward Multiple Regression*

Model		Unstandardized Coefficients		Standardized Coefficients		95.0% Confidence Interval for B		Collinearity Statistics		
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	-.3.06	2.070		-.148	.883	-4.402	3.790		
	English Knowledge	.035	.007	.365	4.725	.000	.021	.050	.896	1.116
	Genre Knowledge	.059	.025	.197	2.302	.023	.008	.109	.730	1.370
	Strategy Use	.037	.018	.168	2.022	.045	.001	.073	.774	1.292
	Chinese Writing Score	.014	.016	.066	.880	.381	-.018	.046	.941	1.062
	Working Memory Span	.022	.031	.053	.706	.481	-.040	.084	.935	1.070

a. Dependent Variable: English Writing Score

Table 4.28 presents the coefficients of each of the independent variables. A comparison of the standardized coefficients revealed that English Knowledge has the greatest impact on English Writing Score (Beta = .365), followed by Genre Knowledge (.197), Strategy Use (.168), Chinese Writing Score (.066) and Working Memory Span (.053). The unstandardized coefficients provide

the information for the regression equation as a whole. Thus the equation for predicting English Writing Score from the five independent variables with the measures used in this study is generated as follows:

$$EW = -.3.06 + .035 EK + .059 GKE + .037 ESU + .014 CW + .022 WM$$

An important step in multiple regression analyses is to verify that the assumptions are met for running the regression procedure, because if the assumptions are not satisfied, there would be risks of drawing unreliable and misleading inferences (Carver & Nash, 2009). As has been pointed out earlier, multicollinearity problem is a concern for multiple regression analyses. An important indicator of multicollinearity is Tolerance. For each independent variable, the tolerance is $1-R^2$, where R^2 is the coefficient of determination for the regression of that variable on all the remaining independent variables (Norušis, 2006). Low values of tolerance indicate high multivariate correlations. The most common test of multicollinearity is Variance inflation factor (VIF), the number of times the variance of the corresponding parameter estimate is increased due to multicollinearity as compared to as it would be if there were no multicollinearity (von Eye & Schuster, 1998). It is the reciprocal of tolerance and is always ≥ 1 . Values of VIF exceeding 10 are often regarded as severe multicollinearity. For the current analysis, large tolerances (close to 1) and small VIFs (see the last two columns in Table 4.28) indicated that multicollinearity should not be a concern for these independent variables.

Another step is to examine the residuals so as to verify the assumptions for normality and homoskedasticity (homogeneity of variance) are met. Two graphs were generated by SPSS—a normal probability plot for assessing the normality assumption and a scatter plot of standardized predicted values (*ZPRED) versus standardized residuals (*ZRESID) for assessing homogeneity of variance. Figure 4.2 shows that the residuals are basically normally distributed (lying along a

45° upward sloping diagonal line) and Figure 4.3 indicates that the independent variables have equal variances (randomly scattered in a horizontal band around a residual value of zero) (Carver & Nash, 2009), which means that the assumptions for normality and homoskedasticity were met with this data set.

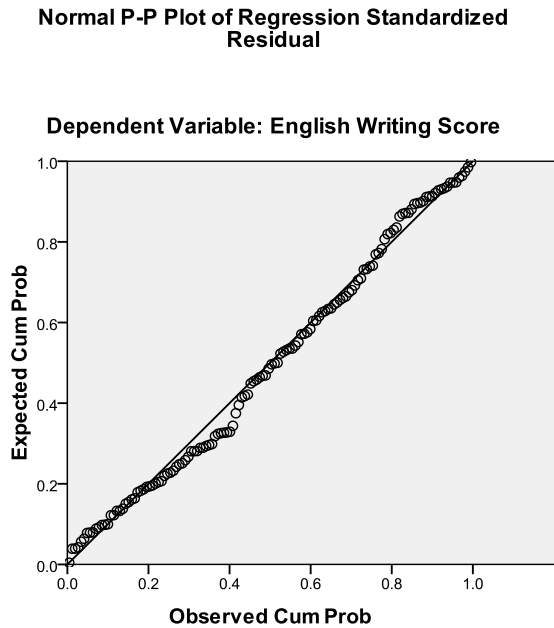


Figure 4.2 Normal Probability Plot of Regression Standardized Residual

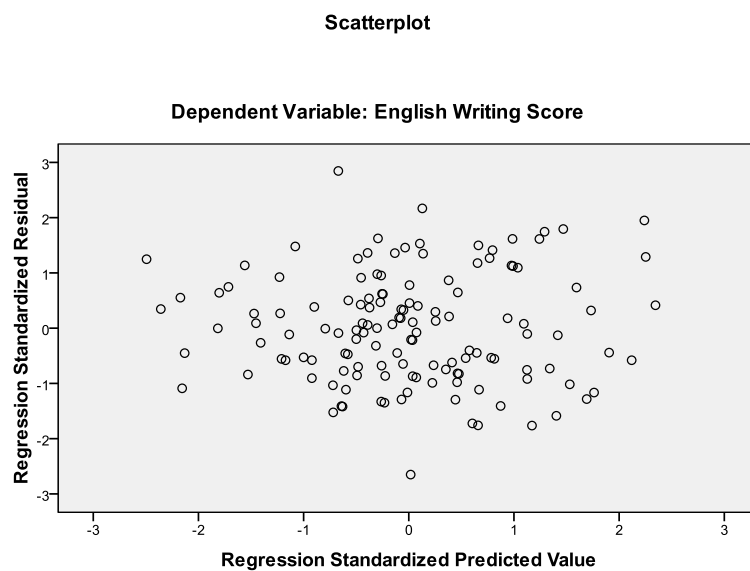


Figure 4.3 Scatter Plot of Residuals

In summary, results from multiple regression analyses show that English Knowledge is the most important predictor of English Writing Score, explaining about 20% of the variance in the dependent variable English Writing Score. The five factors investigated in this study explained about 30% of the variance. However, the small R^2 suggests that the majority of the variance need to be explained by variables other than those included in the model.

CHAPTER 5

DISCUSSION AND CONCLUSION

This chapter summarizes the major findings of the current study and discusses the results in relation to previous studies and literature in relevant fields. Implications of these findings and limitations of this study will also be discussed, and suggestions for future research will be proposed.

Summary of the Findings and Discussion

Research Question 1: Relationship between Writing Ability and Working Memory Capacity

The results showed no correlation between working memory and writing ability, whether in L1 or in L2, which is not consistent with many of the previous studies that investigated similar issues on the relationship between working memory and writing ability (e.g., McCutchen et al., 1994). One of the possible reasons for this discrepancy might be related to the different instruments that were used to measure working memory capacity and writing ability. McCutchen et al. (1994) used short writing tasks for measuring writing ability—primary and secondary students were asked to write 12-15 minutes each on two essay topics in their regular classrooms. Therefore, the operationalization of the construct of writing ability was different in their study from this study, which engaged the participants in two 30-minute academic essay writing tasks in both L1 and L2 in test settings. More importantly, the measures for working memory capacity were very different—McCutchen et al. (1994) employed reading and speaking span tasks for measuring working memory capacity. However, it is self-evident that a person's reading ability is correlated with his/her writing ability (e.g., Carson et al. (1990) reported statistically significant correlations between reading and writing in L1 and L2). In order to avoid this confounding factor of reading ability, operation span tasks rather than reading span tasks were used in the current

study to measure participants' working memory capacity.

Nonetheless, no correlation was found between working memory capacity (using operation span tasks) and writing ability in the current study, which was contrary to what had been expected. This result led to the following questions: in terms of methodology, is working memory capacity task-dependent or task-independent? That is, does an individual's working memory capacity vary with different tasks? Is the operation span task as effective as the reading span task in measuring one's working memory capacity? Is it possible that different mechanisms are involved in performing operation span tasks and reading span tasks?

The choice of operation span tasks in this study was based on the literature which argued for the notion that working memory capacity is domain general (Engle et al, 1992, 1999; Kane et al., 2004) and task independent (La Pointe & Engle, 1990; Turner & Engle, 1989). Engle and his colleagues found in their studies that working memory span tasks tap into the domain general attention control and that working memory capacity is task-independent—operation span tasks are as effective as reading span tasks in measuring one's working memory capacity. However, as Miyake and Shah (1999) have summarized, the limits of working memory capacity reflect multiple factors rather than one single factor. They pointed out that one likely possibility, in terms of the type of tasks, is that “the nature of working memory constraints varies as a function of the novelty and/or complexity of the task” (p. 456).

Even though both Turner and Engle (1989) and La Pointe and Engle (1990) have argued that working memory capacity is independent of the specific nature of the processing component of the span task and that working memory capacity measured with operation span tasks can predict reading comprehension as well as with reading span tasks, it is not clear if the two types of tasks activate or tap into the same storage and processing functions of working memory.

Researchers have proposed different explanations as to what span tasks really measure. For example, Daneman and Carpenter (1980), who proposed the original measure of reading span tasks for working memory capacity, posited that the performance on reading span tasks reflected the amount of resources remaining for storage after processing. However, Engle et al. (1999) proposed that working memory span tasks measure the short-term storage memory capacity plus one's controlled attention ability, particularly the ability of the central executive to coordinate the processing and storage components. In addition, Ericsson and Delaney (1999) argued that individual differences in span tasks reflect differences in encoding and retrieval skills. Therefore, more investigation is needed to find out if operation span tasks and other span tasks are measuring the same working memory capacity. For instance, using both operation and reading span tasks for measuring working memory capacity to replicate the current study (or part of it) with a similar group of Chinese EFL learners and investigating the processes or strategies the participants employ during the span tasks via verbal protocols such as think aloud, retrospective interview, or stimulated recall may reveal if the two types of tasks are measuring the same working memory construct.

Another possible explanation for the lack of correlation between working memory capacity and writing performance might be the inadequacy of the specific instrument used in the current study. The descriptive statistics shows that even though the tasks for measuring working memory capacity achieved acceptably high internal reliability ($\alpha = .824$ for WMC and $.783$ for WME), the item mean was very high ($M = .923$ for WMC and $.855$ for WME) and item variance was rather small ($.067$ for WMC and $.110$ for WME). These items were probably too easy for the participants and thus not differentiating enough. In reality, 18 out of the 136 participants obtained full score in WMC and five in WME. The score distribution was skewed rather than normal,

especially for WMC. There might be a ceiling effect due to the easy span tasks. In fact, Turner and Engle (1989) reported that when the difficulty level of the reading or operation span tasks was moderate, the correlations between working memory span and reading comprehension were higher in magnitude than when the secondary tasks were very simple or very difficult. One way to improve the instrument in future research is to redesign the operation span task and include more difficult items so as to increase the difficulty level of the instrument and in turn improve the statistics for normal distribution and variance of the measure. In addition, only one type of task (operation span task) was used in this study. According to Waters and Caplan (2003), memory span on the basis of a single measure tends to be highly inconsistent. Therefore, using more than one task to assess working memory capacity might yield more reliable measurement of memory spans.

Despite the unexpected results in regard to the relationships between working memory capacity and writing ability, the paired-samples *t*-test did show a significant difference in the participants' working memory capacity in Chinese as compared to English, which is consistent with the results from other studies that used other memory span tasks (such as reading span tasks) for measuring participants' working memory capacity in native and foreign languages (e.g., Service et al., 2003; van den Noort et al., 2006). This result added to the evidence for the notion that the functional working memory capacity in one's native language is larger than that in one's second or foreign language.

Research Question 2: Relationship between L1, L2 Writing Ability and L2 Language Knowledge

In terms of the relationship between L1 and L2 writing ability, the results of this study showed significant but low correlations between English Content scores and Chinese Content and Language scores. However, no significant correlations were found between English

Language scores and Chinese Content or Language scores, which is inconsistent with many studies that reported different degrees of correlation between L1 and L2 writing ability (e.g., Cumming, 1989; Leung, 1984; Sasaki & Hirose, 2002; Schoonen et al., 2003; Yun, 2005). Interestingly, this result corroborates the findings of Carson et al. (1990) and Pennington and So (1993), whose participants were also native speakers of Chinese. These controversies will be discussed in terms of relationship between languages and their differences in orthography, contrastive rhetoric, and threshold hypothesis.

The inconsistency regarding the relationship between L1 and L2 writing scores may be attributed to the different language families of the native languages of the participants and the distance between their L1 and L2. In the current study as well as Carson et al. (1990) and Pennington and So (1993), the participants' first language—Chinese—belongs to the Sino-Tibetan language family, which is extremely different from the Indo-European languages such as French in Cumming (1989) or Dutch in Schoonen et al. (2003), or other isolated languages, such as Japanese in Sasaki and Hirose (1996) and Korean in Yun (2005).¹⁴

In the field of second and third language acquisition, there is a consensus that linguistic distance plays an important role in cross-linguistic transfer. For example, Cenoz (2001) investigated verbal transfer in the oral production of English learners as a third language whose other two languages are Spanish, an Indo-European language, and Basque, a non-Indo-European language. The results show that all the participants presented a stronger influence from Spanish than from Basque, regardless of their first language being Basque or Spanish. This preference for using Spanish as a source language attested to the role that linguistic distance plays in cross-linguistic transfer, because Spanish is closer to English than Basque is. The measure of linguistic distance between Spanish and English is 3, according to West and Graham (2004), who

¹⁴ Some scholars categorize Japanese and Korean as Macro-Altaic languages (see Blench, 2008).

calculated the linguistic distance by coding each language with the number of branches used to connect it to English on the basis of a language family tree developed by Chen, Sokal, and Ruhlen (1995, cited in West & Graham, 2004).¹⁵

In regard to the distance between Chinese and English, the measure is 6 on West and Graham's (2004) list, indicating that Chinese is much farther away from English than Spanish and other Indo-European languages, even some Asian languages such as Japanese and Korean, whose measures are both 4 on the list. Although a simple digital measure can hardly capture the differences between two languages, which differ in phonetics and phonology, morphology and syntax, written form and a myriad of other characteristics, it helps to depict the dissimilarity at a glance. Indeed, Chinese and English are way apart in terms of many linguistic features. For example, according to Yang (2008), Chinese and English differ significantly in terms of "the order of grammatical constituents, the degree of grammatical specificity and the location of grammatical realizations" (p. 462). In Chinese, the primary clause, which is semantically more important than the secondary clause, generally appears after the secondary clause of a complex sentence; whereas in English, complex sentences are usually construed in the pattern of the primary clause followed by the secondary clause. The large linguistic distance between Chinese and English might have hindered the transfer of language skills from L1 to L2.

Another possible reason may lie in the radically different writing systems between the participants' L1 and L2, which may be regarded as another aspect of typological differences between the two languages. Chinese and English writing systems contrast sharply in their visual forms—logographic Chinese characters versus alphabetic English, which may have resulted in different mental representations in the brain and different processing mechanisms. Cognitive

¹⁵ There is no information about the linguistic distance between Basque and English because Basque is not included in their list.

studies involving Chinese-English bilingual adults using neuroimaging techniques such as Event-Related-Potential (ERP) (e.g., Liu & Perfetti, 2003; Tan et al., 2001, 2003) have shown that reading Chinese would activate some brain regions that are responsible for coordinating and integrating visual-spatial analyses of the Chinese characters and thus involves more right hemisphere processing as compared to reading English. It might be inferred that processing Chinese characters and English words involves different mechanisms and these differences may make it harder for linguistic skills in Chinese to be transferred to English. In a study investigating cross-language and writing system relationship in biliteracy acquisition, Wang, Perfetti and Liu (2005) found that orthographic skills in Chinese did not predict reading skills in English. They argued that orthographic skills across English and Chinese may be language-specific. Keung and Ho (2008) obtained the same results—no significant correlation between orthographic skills in Chinese L1 and English L2—with young ESL learners in Hong Kong. They postulated that “the visual-spatial skills acquired in recognizing Chinese characters may not facilitate memory of letter sequences” (p. 110).

One might argue that essay writing, whether in Chinese or in English, involves more than retrieving orthographic information of the words being written down. Even so, retrieval of orthographic information is a necessary and indispensable step in the process of writing. The studies mentioned above all focused on reading processes. It would be interesting and probably revealing to investigate the brain activities during the process of essay writing with the help of advanced technology such as ERP and to compare the differences, if there are any, while writing in different languages such as Chinese and English.

Still another possible explanation for the lack of correlation between Chinese and English language scores may come from the field of contrastive rhetoric. Since Kaplan’s (1966)

seminal article “Cultural Thought Patterns in Intercultural Education,” which marked the beginning of contrastive rhetoric, comparative studies into the written discourse of different culture/language origins have flourished despite criticisms from researchers in other disciplines. Connor (2003) appraised the development in contrastive rhetoric in the past 30 years and affirmed the potential of this framework in the field of ESL and EFL writing research and instruction. The basic notion of contrastive rhetoric entails two hypotheses: 1) each language or culture has its own rhetorical conventions and tendencies, and 2) the linguistic and rhetorical conventions of the first language interfere with writing in the second language (Connor, 1996; Kaplan & Grabe, 2002; Wang, 2007). In the field of SLA, a similar notion was put forward by Odlin (1989), which postulates that cross-linguistic differences in discourse may affect linguistic comprehension as well as production. Even though empirical studies in second language writing yielded no consensus in regard to the facilitative or inhibitive role of L1 writing ability on L2 writing performance, it is conceivable, as Sasaki & Hirose (1996) suggested, that L1 writing ability might not be so powerful in explaining L2 writing ability “when the two languages have different rhetorical conventions” (p. 156).

Further examination of band scores of Chinese and English writing and the differences in English writing scores among the level groups of Chinese writing showed that those who were poor in the Chinese essay writing did not perform well in the English essay writing, whereas those who performed well in English seemed to be good writers in Chinese as well. However, with this result, a seeming controversy arose: earlier I had argued that it would be difficult for Chinese L1 skills to be transferred to English L2 writing, whereas now I am arguing for such a transfer, or at least correlation, among good Chinese L1 writers. I would like to point out that when I was discussing the difficulty in Chinese-English transfer, I was focusing more on the

language skills, but when I am discussing the possibility of transfer of writing skills among good L1 writers to L2 writing, I am focusing more on the cognitive skills that are involved in essay writing. In fact, the Content scores of Chinese essays and English essays did show a certain degree of significant correlations. The cognitive skills involved in organizing ideas and arguments into coherent essays are probably shared or transferable in L1 and L2 writing.

This connection between L1 and L2 writing could be explained by the interdependence hypothesis and the threshold hypothesis proposed by Cummins (1979, 1981) based on his studies in bilingualism. The interdependence hypothesis purports that abilities involving more cognitively demanding tasks (such as literacy and problem-solving) are common across languages and that certain L1 knowledge or skills can be positively transferred in the process of L2 acquisition. The threshold hypothesis assumes that “those aspects of bilingualism which might positively influence cognitive growth are unlikely to come into effect until the child has attained a certain minimum or threshold level of competence in a second language” (Cummins, 1979, p. 229). Therefore, if a language learner has attained only a very low level of competence in the first or second language, linguistic transfer is unlikely to occur. On the basis of these two hypotheses, it seems logical to infer that a student who is a good writer in L1 has a greater potential to be a good writer in L2.

In light of the linguistic threshold hypothesis (Cummins, 1979), the relationship between L1 and L2 writing scores were reexamined in relation to different levels of L2 language knowledge. The results showed a tendency of increase in the correlations between L1 and L2 writing scores with the increase of L2 proficiency levels, though none of the correlations reached the significance level of .05. The changing strength in the correlations suggests that L1 writing ability may interact with L2 language knowledge in their effect on L2 writing performance.

ANOVA results comparing the English writing scores and the Chinese writing scores respectively across the English level groups indicated that students with higher English proficiency performed significantly better in English essay writing.

Further examination was conducted into the relationship between L2 writing ability and L2 language knowledge. The correlations between the two observed variables for L2 writing ability (EC and EL) and each of the three observed variables of L2 language knowledge (GR, VOC1 and VOC2) were found to be statistically significant, though the strength of associations was relatively low. Linear regression analyses with EC and EL as the dependent variable respectively and the three predictors of L2 language knowledge as the independent variables show that the variance explained by the three variables in this study accounted for only 19.1% of the variance of EC and 22.1% of the variance of EL. This percentage is much lower than those reported by Sasaki & Hirose (1996) and Yun (2005), who also investigated the relationship between L2 writing and L2 language knowledge as part of their studies with learners of English in a foreign language setting. This discrepancy might have resulted from the different instruments employed to measure L2 language knowledge. Sasaki and Hirose (1996) used the structure, listening and vocabulary sections of a comprehensive language test for learners of English, and Yun (2005) used a cloze test; however, in the current study, a receptive vocabulary test, a productive vocabulary test and a grammar test were used as manifestation of L2 language knowledge.

Among the three observed variables, productive vocabulary knowledge (VOC2) had the largest explanatory power for both EC and EL, whereas the impacts of grammar (GR) and receptive vocabulary knowledge (VOC1) seemed to be insignificant and negligible. It should be noted that the measure for grammar knowledge in this study was not the traditional grammar test

but a timed grammaticality judgment task which was attempted to test implicit, automatized grammar knowledge. If a different kind of grammar test had been used, the results might have been very different. In addition, these three measures for L2 language knowledge correlated with each other, especially VOC1 and VOC2. As a result, it is difficult to tease out which factor has more impact on the writing ability due to the possible collinearity among these independent variables. Therefore, in the final analysis using multiple regression, these three variables were combined as one latent variable representing L2 language knowledge.

In summary, investigation into the relationship between L1 and L2 writing ability and L2 language knowledge yielded mixed results compared to previous studies. Possible explanations were provided resorting to theories and empirical studies in different fields such as linguistic distances, cognitive science and language processing, contrastive rhetoric, and bilingualism. Further studies are needed in order to get a clearer picture of the complex relationship between these aspects.

Research Questions 3: Relationship between L1, L2 Writing and the Use of Strategies

Correlation coefficients among the observed variables for L1 and L2 writing ability and the use of writing strategies in L1 and L2 writing showed that writing scores correlated with Genre Knowledge and Strategy Use in the same language but not the other language, that is, English Content and Language correlated with English Genre Knowledge and Strategy Use in English essay writing but not with those in Chinese, and vice versa. It is also noted that Genre Knowledge and Strategy Use in Chinese correlated with those in English, though the magnitude was rather small. These results might be understood in the light of the weak version of Cummins' (1979) interdependence hypothesis.

According to Yun (2005), the weak version of the interdependence hypothesis postulates

a correlated factor model, which hypothesizes that L1 writing ability and L2 writing ability are two specific constructs but are correlated, and the strong version postulates one general factor model, which hypothesizes that one first-order general factor (i.e., writing expertise) directly governs all the variables. Yun's confirmatory factor analysis using structural equation modeling (SEM) demonstrated that correlated factor model was far more reasonable than one general factor model in explaining her data. She argued that writing performance in L1 and L2 may not be affected by one common set of abilities (i.e., writing expertise), but by two specific abilities which are interconnected. In the current study, the result that the use of writing strategies correlated with writing scores in the same language but not the other language suggests separate writing expertise rather than one general, underlying competence. However, the correlation between the use of writing strategies in L1 and L2 suggests some sharing of knowledge and skills in L1 and L2 writing. These results corroborate Yun's (2005) findings.

It is also noted that the magnitude of the correlations between English Genre Knowledge and Strategy Use and English Content and Language scores is greater than the magnitude of correlations between those for Chinese. Further examination with paired-samples *t*-test showed significant differences for nine out of the 25 items. The more practice and familiarity of the topic and the type of writing in English and the use of templates for English essays, as reported by the participants in the questionnaire, might have contributed to the higher correlation between genre knowledge and writing scores in English. In reality, this type of argumentative writing task is very similar to the tasks often used in English tests in China, such as the College Entrance Examinations (English subject) and the nation-wide College English Test (CET, which is used in most universities as an exit exam for the subject of English).¹⁶ Students had received explicit

¹⁶ Students in Tsinghua University, the university where the majority of my participants are studying, are not required to take the CET as an exit exam, but they have to take the Tsinghua English Proficiency Test (TEPT-I),

instructions in their English classes and had practiced writing this type of essays to prepare for those tests. However, in terms of training in Chinese writing, the writing task in the College Entrance Examinations (Chinese subject) is usually free prose writing based on a short given title, which is very different from the task used in this study—an argumentative essay based on a prompt. In addition, college students are not required to take any writing class in Chinese or to pass any exit exam like the English subject. Therefore, they received much less training in writing such argumentative essays in Chinese than in English.

Exploration into the Difficulty Strategies and Revision Strategies revealed both similarities and differences in the process of Chinese and English essay writing. A sharp contrast was seen in the number of participants who reported having difficulties in finding appropriate words while writing the essay—the number is almost three times more for the English essay than for the Chinese essay. However, paired-samples *t*-test did not show any significant difference in the strategies used to deal with such difficulties. Most participants, whether writing the English or the Chinese essay, would pause to find an appropriate word instead of leaving it until after the completion of the whole essay, or use a similar word or find another way to express the same idea instead of giving up the idea and writing something else.

For Revision Strategies, the number of participants who reported rereading and revising the English essay is larger than the number for the Chinese essay (97:72), which is in line with many studies which reported higher frequencies of revision episodes in L2 than in L1 writing (see reviews in Roca De Larios et al., 2002 and Silva, 1993). Among those who reread and revised both their Chinese and English essays, paired-samples *t*-test revealed more similarities rather than differences (only 5 out of the 16 items were significantly different), which is also

which is supposed to have higher requirements than the CET. And one of the writing tasks in TEPT-I is the same as the writing task used in this study.

consonant with past studies such as Gaskill (1986) and Hall (1990). In terms of the differences in revision, *t*-test results showed that the participants focused more on local aspects such as grammar, spelling, and punctuation for the English essay, whereas for the Chinese essay, they focused more on global aspects such as content and meaning. This also corroborated the findings in previous studies which reported more skilled writers attended to more global considerations whereas unskilled writers concentrated more on grammatical and linguistic correction (e.g., Uzawa, 1996; Whalen & Menard, 1995; Zamel, 1983). However, as Porte (1996) pointed out, writing context might influence writers' revision strategies. For example, timed writing may add to the anxiety of the writer which would lead to different revision strategies from writing over a period of time with multiple drafts. For the current study, timed essay writing task probably made it hard for the participants to make extensive changes to the overall structure of the essay due to the time constraint and a concern for neat and clear handwriting on the test paper.

In terms of the use of L1 in the writing of L2 essays, most participants reported using both languages while writing the English essay, whereas only a few reported thinking and writing in English only. Similar findings have been reported as to the frequency of L1 use in L2 writing in studies with think-aloud protocols which provided more detailed information into the writing process and strategy use. For example, Qi (1998) found that for his participant, both L1 and L2 were used to initiate ideas. Wang and Wen's (2002) participants also reported that both L1 and L2 were at their disposal when composing in L2, although they seemed to rely more on L1 for generating and organizing ideas and more on L2 for task examination and text generation. Woodall (2002) reported that all his 28 participants code-switched in their L2 writing processes, ranging from a few seconds to several minutes out of a 25-minute writing session, on the basis of which he asserted that "the most salient *qualitative* difference ... is that in L2 writing, two

languages can be at work at the same time” (p.23). It is also interesting to note in Qi’s (1998) study that when an L2 task was relatively cognitively demanding (e.g., essay writing as opposed to letter writing), the participant tended to use more L1 for cognitive processing. Frequent switching to L1 was reported in the protocol when the participant felt the complexity of the task would be too much of a burden for her to process in her weaker language (L2). Switching to the stronger language is therefore seen as a useful strategy “to minimize the load to which the use of a weaker language might otherwise add” (p. 428), which contributes to compensation for working memory limitations.

In summary, investigation into the use of writing strategies in both L1 and L2 yielded mostly similar results and findings as those reported in previous studies on writing processes, many of which used verbal protocols such as think aloud with relatively small number of participants. The results of this study provided statistical evidence as to the differences and similarities of strategy use between L1 and L2 writing.

Research Question 4: Relative Importance of the Selected Variables to L2 Writing Performance

Multiple regression analyses were performed in order to determine statistically the contribution of the investigated variables to Chinese EFL learners’ L2 writing performance in timed essay writing. The results showed that English Knowledge explained 20.4% of the variance of English Writing Score, followed by Genre Knowledge (7.3%), Strategy Use (2.1%), Chinese Writing Score (0.5%) and Working Memory Capacity (0.3%). With respect to the relative importance of the explanatory variables for L2 writing performance in timed essay writing, English Knowledge was found to be by far the most important predictor, more than three times the other variables. This finding about the primary explanatory power of L2 proficiency concurs with some of the previous studies (Pennington & So, 1993; Sasaki & Hirose, 1996; Yun,

2005), which reported the primary role of L2 language knowledge compared with other factors in explaining the quality of L2 writing. But it is inconsistent with some other studies that found a more significant role for L1 writing ability (Cumming, 1989; Schoonen et al., 2003), that showed a similar impact of L1 writing ability and L2 language knowledge (Hirose & Sasaki, 1994), or that emphasized the skill or strategy transfer from L1 writing (Zamel, 1983). However, as has been pointed out earlier in Chapter 2, caution is needed to compare the relative importance of the variables across studies because the measurements for L2 proficiency in different studies are hardly comparable.

Results from multiple regression analyses also showed that, collectively, the investigated variables accounted for only 30.6% of the variance in L2 writing performance, suggesting that the majority of the variance needs to be explained by variables other than those included in this study. Since the current study attempted to investigate the explanatory factors for L2 writing performance from a cognitive perspective, many social cultural factors were not considered. The relatively low explanatory power of the currently investigated variables indicates that cognitive factors can account for only a small aspect of L2 writing and therefore cannot provide a complete picture of what makes a good writer. Additional factors that might have an influence on L2 writing performance include L2 writing instruction, L2 writing experience, test preparation, teacher and students' beliefs about L2 writing, and other motivational or affective factors (e.g., Rinnert & Kobayashi, 2009; Sasaki, 2002, 2004, 2007; Sasaki & Hirose, 1996; Yun, 2005). Sasaki (2009) investigated the changes in EFL students' writing over 3.5 years from a socio-cognitive perspective with the belief that "cognitive aspects of L2 writing are better explained when considered with the social situations that the learners interact with" (p. 52) and found that individual changes in L2 writing abilities were significantly affected by various sociocultural

factors such as study abroad experiences and L2 writing motivation.

In conclusion, the results of the present study showed that the selected cognitive variables explained about 30% of the variances in L2 writing. Among these variables, English Knowledge was found to be the most important factor, followed by Genre Knowledge and Strategy Use; whereas L1 Writing Ability and Working Memory Capacity seemed to have little impact.

Implications

Implications for the Blueprint of an L2 Writer

In terms of L2 writing research, the current study into the cognitive factors of L2 writing integrated research in L2 writing with research in cognitive psychology. Although the selected variables could only explain about one third of the variance in the dependent variable, the results corroborate some findings of previous studies (e.g., Sasaki & Hirose, 1996; Yun, 2005) in that in EFL settings, learners' L2 Language Knowledge is the most important cognitive factor in the prediction of L2 writing performance. This study also revealed that Genre Knowledge and Strategy Use as measured in this study by the post-writing questionnaire play significant roles in the prediction of the participants' L2 writing performance, though these two factors are not as important as L2 Language Knowledge. In addition, the comparison of use of writing strategies provided detailed information of the similarities and differences in terms of what a writer does in the process of L1 and L2 essay writing. Although drawing a detailed blueprint of an L2 writer would require further research, the results of this study contributed a few strokes from a cognitive perspective.

Implications for L2 Writing Instruction

The findings of this study have pedagogical implications for writing instruction for adult EFL learners in an academic setting. Given the primary effect of L2 language knowledge on the

quality of L2 writing and Cummins' (1979) Threshold Hypothesis which postulates that only when one's L2 language proficiency reaches a certain level can a learner transfer the writing skills in L1 to the writing of L2, the improvement of L2 language proficiency should be an indispensable component of L2 writing instruction, especially in the initial stages.

This study also revealed that good L1 writers have the potential to be good L2 writers. On the basis of this finding, coupled with the Interdependence Hypothesis (Cummins, 1979) which posits a common academic and cognitive capability across languages, L2 writing instruction should try to raise the learners' awareness that they can draw on and utilize their knowledge and expertise in L1 writing. More importantly, L2 writing instructors need to be aware of and be able to diagnose the sources of their students' writing problems—from lack of L2 language knowledge, or lack of writing expertise, or both, or some other reasons—and plan their lessons or curriculum accordingly.

The findings about the contribution of Genre Knowledge and Strategy Use suggest that the instruction about genre knowledge and writing strategies would be conducive to the improvement of learners' writing performance. Explicit explanation and training of the writing conventions and the writing strategies pertaining to academic English writing may facilitate students' genre knowledge and strategy use in the writing of academic English essays. In fact, some participants commented in the questionnaire that taking the survey helped them come to know some good strategies that they had never thought of or been told before. Therefore, the use of such questionnaire for writing strategies might be useful for raising students' awareness of the possible writing strategies they can use during the writing process.

Implications for Assessment and/or Measurement

Methodologically, a battery of tests was used to measure the various contributing factors

of L2 writing, some tests borrowed and others constructed. Analyses of reliability statistics showed that some instruments were more reliable than others; however, high reliability does not necessarily entail high validity. For instance, the two vocabulary tests borrowed from Nation and his colleagues turned out to be too difficult for this group of participants whereas the working memory capacity tasks turned out to be too easy. The design or choice of tests has to take into consideration test takers' characteristics. In addition, the measurement of working memory capacity with operation span tasks raised the question about the underlying construct of working memory—whether it is task-dependent or task-independent. More research is in order with other types of instruments/tasks.

Limitations and Suggestions

The current study attempted an interdisciplinary investigation which connected different fields such as second language writing, language assessment and cognitive psychology. The original purpose of the study was to focus on a number of contributing factors for L2 writing ability from a cognitive perspective, especially the role of working memory capacity, as well as their relative importance to the writing performance of Chinese EFL learners in timed essay writing. The results showed that it is far from adequate to focus on only a few cognitive factors because writing is such a complex task. Cognitive explanations, despite their usefulness in accounting for many aspects of writing expertise, do not provide the complete picture of what makes a good writer (Weigle, 2005). Social and cultural factors are indispensable for a learner to become an expert writer within a certain discourse community. As discussed by many researchers in the field of writing (e.g., Grabe, 2001; Grabe & Kaplan, 1996; Hayes, 1996), there is a set of underlying theories that will influence the construct of writing, including a theory of language, of conceptual knowledge and mental representations, of writing processes, of

motivational and affective variables, of social context influences, and of learning. When Ortega (2009) looked ahead for the needed knowledge about EFL writing, she called on that “we will need to understand better the relative importance of cognitive versus affective and attitudinal influences on the knowledge that informs writer choices” (p. 246). Investigation into such affective and attitudinal factors would present new challenges for researchers. Questions such as how should these constructs be defined and measured would be among the first to be answered (Weigle, personal communication, July 7, 2010).

It is also unexpected that this study did not find any significant correlation between working memory capacity and writing performance. As discussed earlier, one of the possible reasons might be the inadequacy of the instruments used for measuring working memory capacity. For future research, the instruments for measuring working memory capacity should be improved, and other instruments such as reading span tasks might be included.

In this study, the participants’ first and second language writing ability was operationalized as their performance (scores assigned by two trained raters) on a timed essay writing. Although timed essay writing is the most direct and most frequently used method to assess one’s writing ability and it has been used in many large-scale tests of language proficiency and in many empirical studies of writing as well, many researchers (e.g., Brown, 2004; Schoonen et al., 2003) maintain that one writing sample with time constraint is not representative enough of a learner’s overall writing ability. Multiple assignments to assess writing proficiency would be more valid for measuring one’s writing ability. However, due to the time constraint and the already large number of tasks in the battery of tests for this study, it was not practical to include more writing tasks. Nevertheless, other non-test materials such as writing assignments for English classes could have been collected as indicators of writing ability.

As a measure to elicit the use of writing strategies during the writing process, post-writing questionnaires were used for data collection. One concern involving the use of questionnaires is the validity of self-report data. Selinger (1983), in particular, argues that self-report data cannot be independently confirmed and should, therefore, be treated with caution, stating that “introspections are conscious verbalizations of what we think we know” (p. 183) and “conscious verbal reports of learners about their own internal device cannot be taken as a direct representation of internal processing” (p. 189). However, in spite of these caveats, Ellis (1994), O’Malley and Chamot (1990), Skehan (1989) and many other researchers support the use of questionnaires as a viable elicitation procedure due to the degree of structure that such instruments afford. As Dörnyei (2003) points out, a well constructed questionnaire is cost-effective and versatile, the latter meaning questionnaires “can be used successfully with a variety of people in a variety of situations targeting a variety of topics” (p. 10). In effect, a large amount of information was gathered from the questionnaires, but only a small amount was analyzed in regard to the research questions of the current study. For further research, more research questions could be posed in regard to the writing processes involved in both L1 and L2 writing. In addition, the participants’ responses to the open-ended questions could also be explored. Ideally, a more complete inventory of writing strategies used by language learners could be constructed and validated.

Methodologically, more powerful statistical techniques such as structural equation modeling (SEM) might be considered for data analysis. This study was designed as an exploratory study into the cognitive factors that contribute to L2 writing ability. Both observed variables and latent variables were investigated via correlation and multiple regression analyses. For future research, analyses using SEM approach could be conducted. As Raykov &

Marcoulides (2000) stated, SEM is “a comprehensive method for the quantification and testing of theories” which takes into account “the measurement error that is ubiquitous in most disciplines” (p. 1). Unlike many classical approaches for multivariate statistical analysis such as regression, which assume no measurement error in the independent variables, SEM provides explicit estimates of these error terms. Schoonen et al. (2009) suggested that using SEM might contribute to valid comparisons of correlations because SEM takes into account differences in reliability of measures. Another advantage of SEM is that both observed and unobserved (i.e., latent) variables can be incorporated into the modeling procedures (Byrne, 2001). SEM has been proven to be a powerful approach in similar studies such as Schoonen (2005), Schoonen et al. (2002), and Yun (2005). For further research, confirmatory analyses may be carried out with the SEM approach in order to find a best fitting model for the current data.

Conclusion

The current study explored the relationship between Chinese EFL learners’ L2 writing ability and five explanatory factors—L1 writing ability, L2 language knowledge, genre knowledge, use of writing strategies, and working memory capacity. Using a battery of tests for the measurement of these variables, the study examined the relative importance of each of the variables in the prediction of L2 writing performance in timed essay writing and provided a comprehensive analysis of the relationships between and among these variables. The results contributed to a better understanding of what makes a good L2 writer from a cognitive perspective. The findings also have implications for L2 writing assessment and instruction.

REFERENCES

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417-423.
- Baddeley, A. D. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, 36, 189-208.
- Bereiter, C. (1980). Development in writing. In L. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 73-93). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bereiter, C., Burtis, P. J., & Scardamalia, M. (1988). Cognitive operations in constructing main points in written composition. *Journal of Memory and Language*, 27(3), 261-278.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Blench, R. (2008). Stratification in the peopling of China: how far does the linguistic evidence match genetics and archaeology? In A. Sanchez-Mazas, R. Blench, M. D. Ross, I. Peiros & M. Lin (eds.), *Past human migrations in East Asia: Matching archaeology, linguistics and genetics* (pp. 105-132). London: Routledge.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains, NY: Pearson Education, Inc.
- Bryman, A. & Cramer, D. (2009). *Quantitative data analysis with SPSS 14, 15 & 16: A guide for social scientists*. London and New York: Routledge.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.
- Caplan, D. & Waters, G. S. (1999). Verbal working memory and sentence comprehension.

Behavioral and Brain Sciences, 22, 77-126.

- Carson, J., Carrell, P., Silberstein, S., Kroll, B., & Kuehn, P. (1990). Reading-writing relationships in first and second language. *TESOL Quarterly*, 24, 245-266.
- Carver, R. H. & Nash, J. G. (2009). *Doing data analysis with SPSS Version 16*. Australia; Belmont, CA: Brooks/Cole Cengage Learning.
- Case, R., Kurland, M. D., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, 33, 386-404.
- Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third Language Acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Ed.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 8-20). Clevedon; Buffalo, NY: Multilingual Matters.
- Chalhoub-Deville, M., & Deville, C. (2005). A look back at and forward to what language testers measure. In E. Hinkel (Ed.), *Handbook of research in second language teaching and research* (pp. 815-831). Mahwah, NJ: Lawrence Erlbaum Associates.
- Chamot, A. U., & O'Malley, J. M. (1994). Language learner and learning strategies. In N. C. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 371-392). San Diego, CA: Academic Press Ltd.
- Chappelle, C. A., Grabe, W., & Berns, M. (1993). *Communicative language proficiency: Definitions and implications for TOEFL 2000* (ETS Internal Report). Princeton, NJ: Educational Testing Service.
- Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second-language writing*. Cambridge: Cambridge University Press.
- Connor, U. (2003). Changing currents in contrastive rhetoric: Implications for teaching and research. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 218-241). Cambridge: Cambridge University Press.
- Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review*, 12(5), 769-786.
- Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning*, 39, 81-141.
- Cumming, A. (2009). Assessing academic writing in foreign and second languages. *Language Teaching*, 42(1), 95-107.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual

- children. *Review of Educational Research*, 49, 222-251.
- Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3-49). Evaluation, dissemination and Assessment Center, California State University, Los Angeles.
- Cummins, J. (1991). Interdependence of first-and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70-89). Cambridge: Cambridge University Press.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- Daneman, M., & Tardif, T. (1987). Working memory and reading skill re-examined. In M. Coltheart (Ed.), *Attention and performance XII* (pp. 491-508). London: Lawrence Erlbaum Associates.
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Durgunoglu, A. Y. (2002). Cross-linguistic transfer in literacy development and implication for language learners. *Annals of Dyslexia*, 52, 189-204.
- Educational Testing Service. (2008). TOEFL iBT Test Independent Writing Rubrics (Scoring Standards). Retrieved March 3, 2009, from <http://www.nafsa3.org/conferences/regional2009/sessions/tues/pruner%20indepwrihd-t2.pdf>
- Educational Testing Service. (n.d.). Writing Topics. Retrieved February 16, 2009, from <http://www.ets.org/Media/Tests/TOEFL/pdf/989563wt.pdf>
- Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking and points of order. *Studies in Second Language Acquisition*, 18, 91-126.
- Ellis, N. C. (2002). Frequency effects and language processing: Investigating formulaic use and input in future expression. *Studies in Second Language Acquisition*, 24, 143-188.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141-172.
- Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory for comprehension: A test of four hypotheses. *Journal of Experimental Psychology*:

- Learning, Memory, and Cognition*, 18(5), 972-992.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General*, 128(3), 309-331.
- Ericsson, K. A. & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp.257-297). Cambridge: Cambridge University Press.
- Flower, L., & Hayes, J. (1980). The dynamics of composing: Making plans and juggling constraints. In L. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31-50). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & computers*, 35(1), 116-124. For DMDX software download: <http://www.u.arizona.edu/~kforster/dmdx/download.htm>
- Freedman, A., Pringle, I., & Yalden, J. (1983). *Learning to write: First language/second language*. New York: Longman.
- Friedman, N. P. & Miyake, A. (2000). Differential roles for spatial and verbal working memory in the comprehension of spatial descriptions. *Journal of Experimental Psychology: General*, 129, 61-83.
- Friedman, N. P., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods*, 37(4), 581-590.
- Gaskill, W. (1986). Revising in Spanish and English as a second language: A process oriented study of composition. Unpublished doctoral dissertation, University of California, Los Angeles.
- Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gillham, B. (2000). *Developing a questionnaire*. London: Continuum.
- Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: An applied linguistic perspective*. London: Longman.
- Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, Mass: Harvard University Press.
- Hale, G., Taylor, G. Bridgeman, B., Carson, J., Kroll, B., and Kantor, R. (1996). *A study of writing tasks assigned in academic degree programs*. (TOEFL Research Report No. 54).

- Princeton, NJ: Educational Testing Service.
- Hall, C. (1990). Managing the complexity of revision among languages. *TESOL Quarterly*, 24(1), 43-60.
- Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14, 25-38.
- Hatasa, Y. A., & Soeda, E. (2000). Writing strategies revisited: A case of non-cognate L2 writers. In B. Swierzbina, F. Morris, M. E. Anderson, C. A. Klee, & E. Tarone (Eds.), *Social and cognitive factors in second language acquisition: Selected proceedings of the 1999 second language research forum* (pp. 375-396). Somerville, MA: Cascadilla.
- Hatch, E. & Lararaton, A. (1991). *The research manual: Design and statistics for applied Linguistics*. New York: Newbury House.
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing* (pp. 1-27). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hayes, J. R., & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg & E. Steinberg (Eds.), *Cognitive process in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Henry, K. (1996). Early L2 writing development: A study of autobiographical essays by university-level students of Russian. *The Modern Language Journal*, 80, 309-326.
- Hirose, K. & Sasaki, M. (1994). Explanatory variables for Japanese students' expository writing in English: An exploratory study. *Journal of Second Language Writing*, 3, 203-229.
- Jacobs, S. (1982). *Composing and coherence: The writing of eleven pre-medical students*. Washington D. C.: Center for Applied Linguistics.
- Jones, S. & Tetroe, J. (1987). Composing in a second language. In A. Matsuhashi (Ed.), *Writing in real time: Modeling production processes* (pp. 34-57). Norwood, N.J.: Ablex Pub. Corp.
- Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, 102(2), 199-225.
- Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, 21(2), 121-151.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122-149.

- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189-217.
- Kaplan, R. B. (1966). Cultural thought patterns in intercultural education. *Language Learning*, *16*, 1-20.
- Kaplan, R. B., & Grabe, W. (2002). A modern history of written discourse analysis. *Journal of Second Language Writing*, *11*(3), 191-223.
- Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57-71). Mahwah, NJ: Lawrence Erlbaum Associates.
- Keung, Y.-C. & Ho, C. S.-H. (2008). Transfer of reading-related cognitive skills in learning to read Chinese (L1) and English (L2) among Chinese elementary school children. *Contemporary Educational Psychology*, *34*, 103-112.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Krapels, A. R. (1990). An overview of second language writing process research. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 37-56). Cambridge: Cambridge University Press.
- Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. New York: McGraw-Hill.
- La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1118-1133.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, *16*, 307-322.
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 33-51.
- Leki, I. (1995). Coping strategies of ESL students in writing tasks across the curriculum. *TESOL Quarterly*, *29*, 235-260.
- Leki, I. (1996). L2 composing: Strategies and perceptions. In B. Leeds (Ed.), *Writing in a second language: Insights from first and second language teaching and research* (pp. 27-36). White Plains, NY: Longman.

- Levy, C. M., & Ransdell, S. (1995). Is writing as difficult as it seems? *Memory and Cognition*, 23, 767-779.
- Liu, Y., & Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: An ERP study of Chinese bilinguals. *Human Brain Mapping*, 18(3), 167-175.
- Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181-209). Philadelphia: John Benjamins.
- Manchon, R. M. (2001). Trends in the conceptualizations of second language composing strategies: A critical analysis. *International Journal of English Studies*, 1(2), 47-70.
- McCutchen, D. (1994). The magical number of three, plus or minus two: Working memory in writing. In E. C. Butterfield (Ed.), *Children's writing: Toward a process theory of the development of skilled writing* (pp. 1-30). Greenwich, CT: JAI.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, 8(3), 299-325.
- McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist*, 35(1), 13-23.
- McCutchen, D., Covill, A., Hoynes, S. H., & Mildes, K. (1994). Individual differences in writing: Implications of translating fluency. *Journal of Educational Psychology*, 86, 256-266.
- McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, 55, 381-401.
- McNamara, T. F. (1996). *Measuring second language performance*. London: Longman.
- Ministry of Education in China. (2009). *2009 Putong Gaodeng Xuexiao Zhaosheng Quanguo Tongyi Kaoshi Dagang (Syllabus for college entrance examinations for matriculation of college students in 2009)*. Beijing, China: Ministry of Education.
- Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339-364). Mahwah, NJ: Lawrence Erlbaum Associates.
- Miyake, A., & Shah, P. (1999). Toward unified theories of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp.442-481). Cambridge: Cambridge University Press.

- Nation, P., & Gu, P. Y. (2007). *Focus on vocabulary*. Sydney, Australia: National Center for English Language Teaching and Research, Macquarie University.
- Norušis, M. J. (2006). *SPSS 15.0 statistical procedures companion*. Upper Saddle River, NJ : Prentice Hall.
- O'Brien, I., Segalowitz, N. S., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, 27, 377-402.
- Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Oller, J. W., Jr. (1979). *Language tests at school: A pragmatic approach*. London: Longman.
- O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Ortega, L. (2009). Studying writing across EFL contexts: Looking back and moving forward. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching, and research* (pp. 232-255). Bristol/Buffalo/Toronto: Multilingual Matters.
- Osaka, M., & Osaka, N. (1992). Language-independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 30, 287-289.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House Publishers.
- Pennington, M. C., & So, S. (1993). Comparing writing process and product across two languages: A study of 6 Singaporean university students' writers. *Journal of Second Language Writing*, 2, 41-63.
- Petrić, B., & Czár, B. (2003). Validating a writing strategy questionnaire. *System*, 31, 187-215.
- Polio, C. (2003). Research on second language writing: An overview of what we investigate and how. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 35-65). Cambridge: Cambridge University Press.
- Porte, G. (1996). When writing fails: How academic context and past learning experiences shape revision. *System*, 24(1), 107-116.
- Purpura, J. E. (1999). *Learner strategy use and performance on language tests: A structural equation modeling approach* (Vol. 8). Cambridge: Cambridge University Press.

- Qi, D. S. (1998). An inquiry into language-switching in second language composing processes. *The Canadian Modern Language Review*, 54(3), 413-435.
- Raimes, A. (1985). What unskilled ESL students do as they write: a classroom study of composing. *TESOL Quarterly*, 19(2), 229-258.
- Raimes, A. (1987). Language proficiency, writing ability, and composing strategies: A study of ESL college student writers. *Language Learning*, 37(3), 439-468.
- Ransdell, S., & Levy, C. M. (1996). Working memory constraints on writing quality and fluency. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 93-105). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ransdell, S., & Levy, C. M. (1999). Writing, reading, and speaking memory spans and the importance of resource flexibility. In M. Torrance & G. Jeffrey (Eds.), *The cognitive demands of writing: Processing capacity and working memory in text production* (pp. 99-113). Amsterdam: Amsterdam University Press.
- Ransdell, S., Arecco, M. R., & Levy, C. M. (2001). Bilingual long-term working memory: The effects of working memory loads on writing quality and fluency. *Applied Psycholinguistics*, 22, 113-128.
- Raykov, T. & Marcoulides, G. A. (2000). *A first course in structural equation modeling*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Reid, J. (1990). Responding to different topic types: A quantitative analysis from a contrastive rhetoric perspective. In B. Kroll (Ed.), *Second language writing: Research insights from the classroom* (pp. 191-210). Cambridge: Cambridge University Press.
- Rinnert, C. & Kobayashi, H. (2009). Situated writing practices in foreign language settings: The role of previous experience and instruction. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching and research* (pp. 23-48). Bristol/Buffalo/Toronto: Multilingual Matters.
- Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on incidental SLA. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 211-251). Philadelphia: John Benjamins.
- Roca de Larios, J., Murphy, L., & Marin, J. (2002). A critical examination of L2 writing process research. In S. Ransdell & M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 11-47). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study.

- Journal of Second Language Writing*, 9(3), 259-291.
- Sasaki, M. (2002). Building an empirically-based model of EFL learners' writing processes. In S. Ransdell & M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 49-80). Dordrecht: Kluwer Academic Publishers.
- Sasaki, M. (2004). A multiple-data analysis of the 3.5-year development of EFL student writers. *Language Learning*, 54(3), 525-582.
- Sasaki, M. (2007). Effects of study-abroad experiences on EFL writers: A multiple-data analysis. *Modern Language Journal*, 91, 602-620.
- Sasaki, M. (2009). Changes in English as a foreign language students' writing over 3.5 years: A sociocognitive account. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching and research* (pp. 49-76). Bristol/Buffalo/Toronto: Multilingual Matters.
- Sasaki, M., & Hirose, K. (1996). Explanatory variables for EFL students' expository writing. *Language Learning*, 46(1), 137-174.
- Scardamalia, M. (1981). How children cope with the cognitive demands of writing. In C. H. Frederiksen & J. F. Dominic (Eds.), *Writing: The nature, development, and teaching of written communication, Vol. 2. Writing: Process, development, and communication* (pp. 81-103). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Schoonen, R. (2005). Generalizability of writing scores: An application of structural equation modeling. *Language Testing*, 22(1), 1-30.
- Schoonen, R., Snellings, P., Stevenson, M. & Van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. In R. M. Manchón (Ed.), *Writing in foreign language contexts: Learning, teaching and research* (pp. 77-101). Bristol/Buffalo/Toronto: Multilingual Matters.
- Schoonen, R., Van Gelderen, A., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., and Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing and metacognitive knowledge. *Language Learning*, 53(1), 165-202.
- Schoonen, R., Van Gelderen, A., De Glopper, K., Hulstijn, J., Snellings, P., Simis, A., and Stevenson, M. (2002). Linguistic knowledge, metacognitive knowledge and retrieval speed in L1, L2 and EFL writing: A structural equation modeling approach. In S. Ransdell & M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 101-122). Dordrecht: Kluwer Academic Publishers.
- Selinger, H. W. (1983). The language learner as linguist: Of metaphors and realities. *Applied Linguistics*, 4, 179-191.

- Service, E., Simola, M., Metsaenheimo, O., & Maury, S. (2002). Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology, 14*, 383-407.
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General, 125*, 4-27.
- Shaw, S. D. & Weir, C. J. (2007). *Examining writing: Research and practice in assessing second language writing*. Cambridge: Cambridge University Press.
- Silva, T. (1990). A comparative study of the composing of selected ESL and native English speaking freshman writers. *Dissertation Abstracts International, 51*(10), 3397A.
- Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing. *TESOL Quarterly, 27*, 657-677.
- Skehan, P. (1989). *Individual differences in second-language learning*. London: Edward Arnold.
- Sparks, R., Ganschow, L., Patton, J., Artzer, M., Siebenhar, D., & Plageman, M. (1997). Prediction of foreign language proficiency. *Journal of Educational Psychology, 89*, 549-561.
- Stevens, J. P. (2007). *Intermediate statistics: A modern approach* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J. H. (2001). The neural system underlying Chinese logographic reading. *NeuroImage, 13*(5), 836-846.
- Tan, L. H., Spinks, J. A., Feng, C. M., Siok, W. T., Perfetti, C. A., Xiong, J. Fox, P., & Gao, J. H. (2003). Neural systems of second language reading are shaped by native language. *Human Brain Mapping, 18*(3), 158-166.
- Tsinghua University. (2009). Introduction to Tsinghua English Proficiency Test-I. Retrieved September 28, 2009, from http://www.tsinghua.edu.cn/docsn/wyx/chinese/TEPT/tept1%20intro_for%20on%20line.htm
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task-dependent? *Journal of Memory and Language, 28*, 127-154.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the

- operation span task. *Behavior Research Methods*, 37, 498-505.
- Uzawa, K. (1996). Second language learners' processes of L1 writing, L2 writing, and translation from L1 into L2. *Journal of Second Language Writing*, 5(3), 271-294.
- Van den Noort, M., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist*, 11(4), 289-296.
- Van den Noort, M., Bosch, P., Haverkort, M., & Hugdahl, K. (2008). A standard computerized version of the reading span test in different languages. *European Journal of Psychological Assessment*, 24(1), 35-42.
- von Eye, A. & Schuster, C. (1998). *Regression analysis for social sciences*. San Diego, CA: Academic Press.
- Wang, M., Perfetti, C. A., & Liu, Y. (2005). Chinese-English biliteracy acquisition: Cross-language and writing system transfer. *Cognition*, 97, 67-88.
- Wang, W. (2007). Contrastive rhetoric and studies of Chinese and English writing. Retrieved April 14, 2010, from <http://ses.library.usyd.edu.au/bitstream/2123/1701/3/03chapter2.pdf>
- Wang, W.-Y., & Wen, Q.-F. (2002). L1 use in the L2 composing process: An exploratory study of 16 Chinese EFL writers. *Journal of Second Language Writing*, 11, 225-246.
- Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology*, 49A(1), 51-74.
- Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*, 35(4), 550-564.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Weigle, S. C. (2005). Second language writing expertise. In K. Johnson (Ed.), *Expertise in second language learning and teaching* (pp. 128-149). Basingstoke, Hampshire/ New York, NY: Palgrave Macmillan.
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Basingstoke, England/ New York, NY: Palgrave Macmillan.
- West, J. & Graham, J. L. (2004). A linguistic-based measure of cultural distance and its relationship to managerial values. *Management International Review*, 44, 239-260.
- Whalen, K., & Menard, N. (1995). L1 and L2 writers' strategic and linguistic knowledge: A model of multiple-level discourse processing. *Language Learning*, 45(3), 381-418.

- White, E. M. (1984). Holisticism. *College Composition and Communication*, 35(4), 400-409.
- Williams, J. N. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition*, 21, 1-48.
- Woodall, B. R. (2002). Language-switching: Using the first language while writing in a second language. *Journal of Second Language Writing*, 11, 7-28.
- Yang, Y. (2008). Typological interpretation of differences between Chinese and English in grammatical metaphor. *Language Sciences*, 30, 450-478.
- Yun, Y. (2005). *Factors explaining EFL learners' performance in a timed essay writing test: A structural equation modeling approach*. Unpublished Doctoral dissertation, University of Illinois at Urbana-Champaign, Urbana, Illinois.
- Zamel, V. (1983). The composing processes of advanced ESL students: Six cases studies. *TESOL Quarterly*, 17, 165-187.

APPENDIX A
ESSAY WRITING PROMPTS

English Writing Task

Instructions:

Please write an essay in English on the following topic. You will have 30 minutes to complete the essay. Please write no less than 150 words. If you need more space, please turn over to continue.

Prompt:

Some people choose their major field of study based on their personal interests, while others are more concerned about future employment possibilities. What position do you support? Use specific reasons and examples to support your answer.

Chinese Writing Task

Instructions:

Please write an essay in English on the following topic. You will have 30 minutes to complete the essay. Please write no less than 300 words. If you need more space, please turn over to continue.

Prompt:

在填报志愿选择大学时，你考虑得更多的是你想学习的专业还是你想上的大学的名声和地位，或是其他某个对你来说更重要的因素？为什么？请使用具体的理由和事例来支持你的观点。(Which factor do you consider most important in your decision when applying for a college/university, the major you are interested in, the reputation of the college/university, or any other factor that is more important for you? Use specific reasons and examples to support your opinion.)

APPENDIX B
QUESTIONNAIRES

Research ID _____

Background Information

1. Name: _____
2. E-mail address: _____
3. Gender: Male _____ Female _____
4. Age: _____
5. University: _____
6. Major: _____
7. Are you an undergraduate or a post-graduate? UG (本科) _____ G (研究生) _____
8. Which year are you in?
 Freshman (1st year) _____ Sophomore (2nd year) _____
 Junior (3rd year) _____ Senior (4th year) _____
9. How old were you when you started to learn English? _____
10. How long have you been studying English? _____
11. What was your score for English in the College Entrance Examination? _____
12. Have you taken CET-4 and/or CET-6?
 CET-4 No _____ Yes _____ → When _____ Score _____
 CET-6 No _____ Yes _____ → When _____ Score _____
13. Have you taken TOEFL? If yes, what is your best score?
 No _____
 Yes _____ → When _____ Best score _____/TWE score _____

Please circle one of the choices for the following questions:

14. How do you rate your overall proficiency in English?
 Excellent Good Fair Poor
15. How do you rate your writing ability in English?
 Excellent Good Fair Poor
16. How do you rate your writing ability in Chinese?
 Excellent Good Fair Poor
17. How much English writing (including essays, letters, email, messages, journals/diary, etc.) do you do in a week?
 < 1 hour 1-3 hours 3-5 hours > 5 hours
18. How much Chinese writing (including essays, letters, email, messages, journals/diary, etc.) do you do in a week?
 < 1 hour 1-3 hours 3-5 hours > 5 hours

Research ID: _____

Post-writing Questionnaire

(English version)

This questionnaire is designed to gather information about how you dealt with the English writing task. There are 56 statements about what you might have done when you wrote the English essay. Please circle the corresponding number to indicate how true each statement is in terms of **what you actually did**. There are no “right” or “wrong” answers. The data we collect are for research purposes and your opinions will be kept confidential. Thank you very much for your cooperation!

Scales:

1= strongly disagree

2= disagree

3= not sure

4= agree

5= strongly agree

0= not applicable (N/A)

Items:

	strongly disagree				strongly agree	N/A
1. I'm familiar with the type of essays I just wrote (i.e., argumentative essay).	1	2	3	4	5	0
2. I know how to write this type of essays.	1	2	3	4	5	0
3. I have practiced writing this type of essays.	1	2	3	4	5	0
4. I had a ready framework/template for this type of essays.	1	2	3	4	5	0
5. I used a ready framework/template in my writing.	1	2	3	4	5	0
6. I'm familiar with the topic I just wrote about.	1	2	3	4	5	0
7. I liked writing about this topic.	1	2	3	4	5	0
8. I had enough ideas to write about this topic.	1	2	3	4	5	0
9. I have practiced writing this topic or a similar topic.	1	2	3	4	5	0
10. Before I started writing, I read the prompt carefully to make sure what was required.	1	2	3	4	5	0
11. I thought of what content I should include before I started writing.	1	2	3	4	5	0
12. I thought of how to organize my ideas before I started writing.	1	2	3	4	5	0

13. I thought of how to satisfy readers or examiners. 1 2 3 4 5 0
14. Before I started writing, I wrote down some notes such as main arguments and keywords. 1 2 3 4 5 0
15. Before I started writing, I made an outline on paper or in my mind. 1 2 3 4 5 0
16. I started writing the essay without a written or mental plan. 1 2 3 4 5 0

What else did you do before writing?

17. I stated my main points in the first paragraph. 1 2 3 4 5 0
18. I wrote at least three paragraphs. 1 2 3 4 5 0
19. I used examples to support my ideas. 1 2 3 4 5 0
20. I ended my essay with a conclusion. 1 2 3 4 5 0
21. I felt it was easy to put ideas in good order. 1 2 3 4 5 0
22. I felt it was easy to express ideas using appropriate words. 1 2 3 4 5 0
23. I felt it was easy to express ideas using correct sentences/grammar. 1 2 3 4 5 0
24. I was concerned about grammatical accuracy when I was writing. 1 2 3 4 5 0
25. I tried to connect my idea smoothly when I was writing. 1 2 3 4 5 0
26. I sometimes had difficulties when I tried to express my ideas with appropriate words.
- a) Yes b) No

If you answered No to Question 26, skip Questions 27-33.

27. When I had a problem expressing my ideas, I stopped and thought for a while. 1 2 3 4 5 0
28. I continued writing and came back to the problem later. 1 2 3 4 5 0
29. I used a similar (English) word that I could think of. 1 2 3 4 5 0
30. I tried to think of another way to express my ideas. 1 2 3 4 5 0

- | | | | | | | |
|--|---|---|---|---|---|---|
| 31. I gave up and decided to write something else. | 1 | 2 | 3 | 4 | 5 | 0 |
| 32. I only used words which I was sure were correct. | 1 | 2 | 3 | 4 | 5 | 0 |
| 33. *I used Chinese and later tried to find an appropriate English word. | 1 | 2 | 3 | 4 | 5 | 0 |

What else did you do during writing?

34. I spent some time reading back and revising what I had written down.

- a) Yes b) No

If you answered No to Question 34, skip Questions 35-50.

- | | | | | | | |
|--|---|---|---|---|---|---|
| 35. I often read back during the whole writing process. | 1 | 2 | 3 | 4 | 5 | 0 |
| 36. I read what I had written down after I finished the whole essay. | 1 | 2 | 3 | 4 | 5 | 0 |
| 37. I revised as I wrote throughout the writing of the entire essay. | 1 | 2 | 3 | 4 | 5 | 0 |
| 38. I started revising after I finished writing the entire essay. | 1 | 2 | 3 | 4 | 5 | 0 |

When I was revising,

- | | | | | | | |
|--|---|---|---|---|---|---|
| 39. I focused mostly on content and meaning. | 1 | 2 | 3 | 4 | 5 | 0 |
| 40. I focused mostly on organization. | 1 | 2 | 3 | 4 | 5 | 0 |
| 41. I focused mostly on grammar, spelling, and punctuation. | 1 | 2 | 3 | 4 | 5 | 0 |
| 42. I focused on both content and grammar. | 1 | 2 | 3 | 4 | 5 | 0 |
| 43. I corrected spelling and/or grammatical mistakes. | 1 | 2 | 3 | 4 | 5 | 0 |
| 44. I changed my wording to express my ideas more clearly. | 1 | 2 | 3 | 4 | 5 | 0 |
| 45. I made changes in sentence structure where necessary. | 1 | 2 | 3 | 4 | 5 | 0 |
| 46. I made changes in the organization of the essay where necessary. | 1 | 2 | 3 | 4 | 5 | 0 |
| 47. I made changes in the content of the essay where necessary. | 1 | 2 | 3 | 4 | 5 | 0 |
| 48. I tried to make my writing more coherent. | 1 | 2 | 3 | 4 | 5 | 0 |
| 49. I checked if my writing matched the requirements. | 1 | 2 | 3 | 4 | 5 | 0 |

50. I thought of if my writing would satisfy reader/examiner. 1 2 3 4 5 0

What else did you do for revising?

51. *I formulated my ideas mostly in Chinese
and translated them into English. 1 2 3 4 5 0

52. *I generated ideas mainly in Chinese,
but formulated sentences in English. 1 2 3 4 5 0

53. *I generated ideas in both languages,
but formulated sentences in English. 1 2 3 4 5 0

54. *I thought mostly in English and wrote in English. 1 2 3 4 5 0

Please write any of your other comments here:

Note: Five items with a * at the beginning (#32, #51-54) ask about the use of L1 (Chinese) in L2 (English) essay writing, therefore they only appear in the questionnaire for L2 essay writing. They are not included in the questionnaire for Chinese essay writing.

Post-writing Questionnaire

Research ID _____

问卷调查（英文写作）

本问卷旨在调研您在写作过程中所采用的一些策略和措施。问卷中共有 54 个陈述句，**请根据您在刚才的写作过程中所做的一切，圈出相应的数字级别**。答案没有正误之分。本问卷收集的数据仅用于我们的研究，对于您的回答我们将严格保密。谢谢您的参与与合作！

级别

- 1=完全不同意
- 2=不同意
- 3=不确定
- 4=同意
- 5=完全同意
- 0=不适用（N/A）

	完全不同意	→	完全同意	N/A		
1. 我对这种议论文文体很熟悉。	1	2	3	4	5	0
2. 我知道怎么写这种议论文。	1	2	3	4	5	0
3. 我练习过写作这种文体。	1	2	3	4	5	0
4. 我头脑里有议论文文体的框架结构/模板。	1	2	3	4	5	0
5. 我在写作过程中套用了现成的模板。	1	2	3	4	5	0
6. 我对刚刚写作的这个话题很熟悉。	1	2	3	4	5	0
7. 我喜欢这个写作题目。	1	2	3	4	5	0
8. 我有足够的素材和论点来写这个话题。	1	2	3	4	5	0
9. 我练习过写作类似的话题。	1	2	3	4	5	0
10. 动笔之前我仔细地审题并看清了题目的要求。	1	2	3	4	5	0
11. 动笔之前我思考了应该包括哪些内容。	1	2	3	4	5	0
12. 动笔之前我思考了如何来组织我的观点。	1	2	3	4	5	0
13. 动笔之前我思考了如何令阅卷老师感到满意。	1	2	3	4	5	0
14. 写作文之前我做了一些笔记，记下要包括的观点/关键词等。	1	2	3	4	5	0
15. 写作文之前我列了个提纲或打了腹稿（组织了思路）。	1	2	3	4	5	0
16. 我没有列提纲也没有打腹稿，直接就在答题纸上开始写。	1	2	3	4	5	0

开始写作文之前您还做了些什么？（请列举上面没有提及的事项）

- | | | | | | | |
|---------------------|---|---|---|---|---|---|
| 17. 我在第一段简述了我的主要观点。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 18. 我至少写了三段。 | 1 | 2 | 3 | 4 | 5 | 0 |

- | | | | | | | | |
|-----|-----------------------------|---|---|---|---|---|---|
| 19. | 我用具体的事例来支持我的观点。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 20. | 我以结论收尾。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 21. | 我觉得组织论点很容易。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 22. | 我觉得选择恰当的词语来遣词造句很容易。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 23. | 我觉得写出语法正确的句子很容易。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 24. | 写的时候我就已经注意语句是否连贯通顺。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 25. | 写的时候我就已经注意语法是否正确。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 26. | 有时候我找不到恰当的词语来表达我的观点。
是 否 | | | | | | |

如果您对上一题（第26题）的回答是“否”，请跳过27—33题。

- | | | | | | | | |
|-----|-------------------------|---|---|---|---|---|---|
| 27. | 当我词不达意时，我停下来寻找合适的词。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 28. | 我继续写，写完了再回头思考，寻找合适的词。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 29. | 我用一个我能够想到的同义词。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 30. | 我尽量想出别的方法/句子来表达同样的意思。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 31. | 我放弃这个论点，写点别的东西。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 32. | 我只用我认为是准确无误的词。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 33. | *我先写下中文词，然后再用合适的英文词来替代。 | 1 | 2 | 3 | 4 | 5 | 0 |

写作过程中您还做了些什么？（请列举上面没有提及的事项）

34. 我花了一些时间重读并修改我写下的内容。
是 否

如果您对上一题（第34题）的回答是“否”，请跳过35—50题。

- | | | | | | | | |
|-----|-------------------------|---|---|---|---|---|---|
| 35. | 整个写作过程中我经常回头重读已写下的内容。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 36. | 我在写完全文后才通读。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 37. | 整个写作过程中我边写边改。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 38. | 我在写完全文后再开始修改。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 39. | 修改时，我主要关注文章的内容和意思的表达。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 40. | 修改时，我主要关注文章的结构。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 41. | 修改时，我主要关注语法、错别字（拼写）和标点。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 42. | 修改时，我既注重语法，又注重内容。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 43. | 修改时，我更正语法和错别字（拼写错误）。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 44. | 我更改用词，以便更恰当地表达我的思想。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 45. | 在必要的地方，我对句子结构进行了修改。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 46. | 在必要的地方，我对文章的结构进行了修改。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 47. | 在必要的地方，我对文章的内容进行了修改。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 48. | 修改时，我尽力使文章连贯通顺。 | 1 | 2 | 3 | 4 | 5 | 0 |
| 49. | 我核对了我的文章是否符合题目要求。 | 1 | 2 | 3 | 4 | 5 | 0 |

50. 我想到了我的文章能否令阅卷老师感到满意。 1 2 3 4 5 0

修改作文时您还做了什么？（请列举上面没有提及的事项）

51. *我大多数时候用中文构思并造句，然后翻译成英文。 1 2 3 4 5 0

52. *我主要用中文构思，但用英文遣词造句。 1 2 3 4 5 0

53. *我在构思时中英文并用，但用英文遣词造句。 1 2 3 4 5 0

54. *我主要用英文构思并遣词造句。 1 2 3 4 5 0

如果您对本问卷或本研究有任何想法或评论，请在此留言：

非常感谢您的参与！

APPENDIX C

OPERATION SPAN TASKS FOR WORKING MEMORY CAPACITY

Materials for Working Memory Capacity in English:

Stimulus Equations	Yes/No	Words for Recall in L2
IS $(10 \times 1) - 7 = 3$ correct?	y	baby
IS $(10 / 1) + 1 = 10$ correct?	n	cloud
IS $(9 \times 3) + 2 = 27$ correct?	n	mouse
IS $(10 \times 2) - 1 = 19$ correct?	y	sand
IS $(4 / 1) - 3 = 1$ correct?	y	road
IS $(5 \times 2) + 2 = 12$ correct?	y	table
IS $(8 \times 1) + 2 = 10$ correct?	y	hair
IS $(7 \times 1) + 6 = 13$ correct?	y	floor
IS $(7 / 7) + 5 = 6$ correct?	y	lake
IS $(10 / 2) + 4 = 3$ correct?	n	time
IS $(9 / 3) - 2 = 1$ correct?	y	state
IS $(4 / 1) + 1 = 4$ correct?	n	game
IS $(7 \times 2) - 1 = 14$ correct?	n	mind
IS $(2 \times 3) + 1 = 7$ correct?	y	fact
IS $(6 \times 2) + 3 = 15$ correct?	y	word
IS $(8 / 2) + 4 = 8$ correct?	y	sofa
IS $(3 \times 7) + 2 = 24$ correct?	n	group
IS $(9 \times 2) - 4 = 14$ correct?	y	space
IS $(12 / 6) - 3 = 1$ correct?	n	girl
IS $(4 \times 5) - 7 = 11$ correct?	n	desk
IS $(4 / 2) + 1 = 6$ correct?	n	hill
IS $(6 / 2) - 1 = 1$ correct?	n	paper
IS $(9 / 1) + 8 = 18$ correct?	n	class
IS $(6 / 2) - 2 = 2$ correct?	n	form
IS $(8 \times 4) + 2 = 34$ correct?	y	line
IS $(6 \times 2) - 2 = 10$ correct?	y	park
IS $(7 \times 7) + 1 = 49$ correct?	n	east
IS $(8 / 4) + 6 = 8$ correct?	y	music
IS $(3 / 1) + 3 = 5$ correct?	n	check
IS $(7 / 1) + 2 = 7$ correct?	n	ears
IS $(6 / 6) + 2 = 4$ correct?	n	movie
IS $(10 \times 1) - 5 = 10$ correct?	n	page
IS $(5 \times 1) - 1 = 4$ correct?	y	side
IS $(2 \times 1) - 1 = 1$ correct?	y	heart
IS $(9 / 3) + 3 = 6$ correct?	y	town
IS $(10 / 1) + 3 = 13$ correct?	y	river
IS $(10 \times 2) + 2 = 21$ correct?	n	head
IS $(9 \times 3) - 2 = 25$ correct?	y	rain

IS $(2 / 1) - 1 = 1$ correct?	y	lamp
IS $(9 \times 1) - 5 = 5$ correct?	n	world
IS $(10 / 2) + 4 = 9$ correct?	y	face
IS $(3 \times 2) + 1 = 6$ correct?	n	play
IS $(5 / 5) + 4 = 5$ correct?	y	earth
IS $(2 \times 4) + 1 = 8$ correct?	n	night
IS $(8 / 8) + 1 = 2$ correct?	y	come
IS $(10 / 5) - 1 = 1$ correct?	y	store
IS $(10 / 1) - 5 = 4$ correct?	n	lunch
IS $(4 \times 4) + 1 = 17$ correct?	y	grass
IS $(7 \times 2) - 5 = 9$ correct?	y	party
IS $(12 / 2) + 3 = 8$ correct?	n	home
IS $(3 \times 5) + 4 = 18$ correct?	n	deal
IS $(2 \times 4) - 2 = 4$ correct?	n	money
IS $(16 / 4) - 3 = 1$ correct?	y	week
IS $(3 \times 3) + 6 = 15$ correct?	y	apple
IS $(5 \times 2) - 3 = 9$ correct?	n	tree
IS $(14 / 2) + 4 = 11$ correct?	y	ants
IS $(3 \times 6) + 2 = 20$ correct?	y	chair
IS $(3 \times 3) - 2 = 6$ correct?	n	dance
IS $(12 / 3) - 3 = 1$ correct?	y	email
IS $(4 \times 2) + 2 = 12$ correct?	n	show

Materials for Working Memory Capacity in Chinese:

Stimulus Equations	Yes/No	Words for Recall in L1
$(4 \times 2) - 5 = 3$ 是否正确?	y	大学
$(6 / 2) + 1 = 7$ 是否正确?	n	天空
$(8 \times 1) + 9 = 17$ 是否正确?	y	复印
$(10 / 2) + 1 = 6$ 是否正确?	y	机会
$(4 \times 1) - 3 = 1$ 是否正确?	y	孩子
$(6 \times 2) - 2 = 12$ 是否正确?	n	吃饭
$(8 / 1) + 3 = 10$ 是否正确?	n	颜色
$(5 \times 2) - 6 = 4$ 是否正确?	y	地图
$(7 / 7) + 5 = 6$ 是否正确?	y	天气
$(10 \times 2) - 4 = 13$ 是否正确?	n	老师
$(6 / 2) - 3 = 0$ 是否正确?	y	头发
$(4 \times 3) + 1 = 15$ 是否正确?	n	思考
$(8 \times 2) - 3 = 12$ 是否正确?	n	电脑
$(5 \times 3) + 1 = 14$ 是否正确?	n	中国
$(12 / 2) + 3 = 9$ 是否正确?	y	房间
$(6 / 3) - 1 = 1$ 是否正确 ?	y	铅笔
$(9 \times 1) + 8 = 18$ 是否正确?	n	旅游

$(8 \div 2) - 4 = 2$ 是否正确?	n	手机
$(7 \times 4) + 2 = 30$ 是否正确?	y	书包
$(6 \times 2) - 2 = 10$ 是否正确?	y	照片
$(7 \times 7) + 1 = 48$ 是否正确?	n	公园
$(8 \div 4) + 5 = 7$ 是否正确?	y	上课
$(3 \div 1) + 9 = 12$ 是否正确?	y	新闻
$(7 \times 2) + 2 = 15$ 是否正确?	n	冰箱
$(6 \div 6) + 8 = 10$ 是否正确?	n	医生
$(10 \times 1) - 5 = 4$ 是否正确?	n	玩具
$(5 \div 1) - 1 = 4$ 是否正确?	y	手指
$(2 \times 9) + 3 = 21$ 是否正确?	y	调查
$(9 \div 3) + 5 = 6$ 是否正确?	n	学生
$(10 \div 1) - 3 = 13$ 是否正确?	n	电视
$(6 \times 2) + 7 = 19$ 是否正确?	y	家庭
$(8 \times 3) - 2 = 20$ 是否正确?	n	英语
$(12 \div 3) - 3 = 1$ 是否正确 ?	y	游戏
$(5 \times 2) - 3 = 5$ 是否正确?	n	城市
$(10 \div 2) + 4 = 9$ 是否正确 ?	y	花草
$(3 \times 6) - 2 = 14$ 是否正确?	n	黄金
$(5 \div 5) + 5 = 6$ 是否正确 ?	y	耳朵
$(2 \times 4) + 3 = 9$ 是否正确?	n	白领
$(8 \div 8) + 6 = 7$ 是否正确?	y	鼠标
$(10 \div 5) - 2 = 0$ 是否正确?	y	夏天
$(12 \div 2) - 5 = 4$ 是否正确 ?	n	汽车
$(4 \times 4) + 6 = 20$ 是否正确?	n	报纸
$(5 \times 3) - 2 = 13$ 是否正确?	y	语文
$(9 \div 3) - 3 = 1$ 是否正确 ?	n	跑步
$(4 \times 5) - 8 = 15$ 是否正确?	n	挂历
$(10 \div 2) + 6 = 9$ 是否正确 ?	n	公司
$(2 \times 6) - 4 = 8$ 是否正确?	y	绿茶
$(7 \times 3) + 2 = 23$ 是否正确?	y	邮件
$(15 \div 3) - 3 = 1$ 是否正确 ?	n	教授
$(4 \times 3) + 2 = 15$ 是否正确?	n	长江
$(8 \div 2) + 5 = 9$ 是否正确 ?	y	打球
$(3 \times 3) - 2 = 7$ 是否正确?	y	字典
$(2 \times 4) - 5 = 3$ 是否正确?	y	超市
$(6 \div 3) + 4 = 7$ 是否正确 ?	n	工作
$(7 \times 2) - 5 = 9$ 是否正确?	y	黄河
$(12 \div 2) + 4 = 8$ 是否正确 ?	n	考试
$(3 \times 8) + 2 = 26$ 是否正确?	y	电影

$(2 \times 3) - 2 = 4$ 是否正确?	y	北京
$(12 / 4) - 3 = 1$ 是否正确 ?	n	书桌
$(3 \times 3) + 5 = 13$ 是否正确?	n	国庆

Materials for Practice:

Stimulus Equations	Yes/No	Words for Recall in L2	Words for Recall in L1
IS $(9 / 1) + 4 = 14$ correct?	n	phone	地球
IS $(6 / 2) - 2 = 1$ correct?	y	math	宿舍
IS $(9 \times 1) + 9 = 16$ correct?	n	sing	眼睛
IS $(8 \times 1) + 5 = 13$ correct?	y	point	唱歌
IS $(6 \times 2) - 3 = 10$ correct?	n	desk	成功
IS $(8 / 4) - 1 = 1$ correct?	y	read	海洋
IS $(3 / 3) + 5 = 7$ correct?	n	bear	台灯
IS $(10 / 1) + 2 = 12$ correct?	y	child	运动
IS $(10 / 1) - 1 = 9$ correct?	y	watch	数学
IS $(6 \times 1) - 4 = 1$ correct?	n	hole	树木
IS $(10 / 2) - 3 = 2$ correct?	y	room	秘密
IS $(5 \times 2) - 5 = 4$ correct?	n	move	石油
IS $(2 \times 1) - 3 = 6$ correct?	n	book	面包
IS $(7 / 1) - 2 = 5$ correct?	y	door	河流

APPENDIX D
VOCABULARY SIZE TESTS

A Receptive Vocabulary Size Test
(From Nation and Gu, 2007)

Choose the best meaning for each word. If you do not know the word at all, do not guess. Wrong guesses will be taken away from your correct answers. However, if you think you might know the meaning or part of it, then you should try to find that answer.

Thousand 1

1. past: It happened in the <past>.
 - a. time before now
 - b. time of fighting
 - c. time when it is dark
 - d. time of the year when it is hot
2. employ: She was <employed> there.
 - a. married
 - b. a part of a group
 - c. made to stay
 - d. working
3. actual: The <actual> one is larger.
 - a. real
 - b. old
 - c. round
 - d. other
4. shoe: Where is <your shoe>?
 - a. the person who looks after you
 - b. the thing you keep your money in
 - c. the thing you use for writing
 - d. the thing you wear on your foot
5. rid: You could get <rid> of it.
 - a. more
 - b. free
 - c. tired
 - d. less
6. admissible: That is not <admissible>.
 - a. able to be believed
 - b. allowed
 - c. able to be described
 - d. approved
7. finery: She was dressed in her <finery>.
 - a. loose, light coat worn in the house by women
 - b. delicate underclothes
 - c. most expensive clothes and jewelry
 - d. piece of animal fur to cover the shoulders
8. army: They saw the <army>.
 - a. black and white animal
 - b. place for keeping books
 - c. person from the next house

- d. group of fighting men
- 9. behavior: Look at <her behavior>!
 - a. the people who have come to listen to her
 - b. the way she acts
 - c. the large amount of money she has
 - d. the land she owns
- 10. refuse: She <refused>.
 - a. said no
 - b. thought about it
 - c. returned
 - d. stayed late

Thousand 2

- 1. mortgage: We need a large <mortgage>.
 - a. loan
 - b. place for watching games
 - c. piece of land round a house
 - d. group of people to work for us
- 2. lend: She often <lends her books>.
 - a. gives her books to someone else for a time
 - b. writes on the pages of her books in a bad way
 - c. cleans her books
 - d. puts her name in her books
- 3. delay: He was <delayed>.
 - a. hurt
 - b. attacked suddenly
 - c. made angry
 - d. made late
- 4. hire: He <hired the boat>.
 - a. paid to use the boat
 - b. turned the boat upside down in the water
 - c. fixed the boat
 - d. cleaned the bottom of the boat
- 5. upset: I am <upset>.
 - a. tired
 - b. famous
 - c. rich
 - d. unhappy
- 6. super: That was <super>!
 - a. excellent
 - b. very uncomfortable
 - c. extremely important
 - d. stupid
- 7. pin: I need a <pin>.
 - a. special piece of glass that makes things look bigger or smaller
 - b. metal stick with a sharp point at one end
 - c. thing with ink for writing

- d. short holiday
- 8. patience: He <has no patience>.
 - a. will not wait happily
 - b. has no free time
 - c. has no faith
 - d. does not know what is fair
- 9. nil: His mark for that question was <nil>.
 - a. very bad
 - b. nothing
 - c. very good
 - d. in the middle
- 10. marvel: We <marveled> at it.
 - a. looked carefully
 - b. shouted loudly
 - c. were filled with surprise
 - d. were very disappointed

Thousand 3

- 1. steer: Let me <steer> it.
 - a. look carefully in and around it
 - b. make it better
 - c. put oil onto it
 - d. make it go where I want
- 2. cure: Can you <cure it>?
 - a. touch it gently
 - b. understand it
 - c. explain it in more detail
 - d. make it well again
- 3. mug: This <mug> needs a wash.
 - a. tall cup without a saucer
 - b. old car that you are fond of
 - c. piece of clothing worn next to the skin
 - d. sheltered place in front of a door
- 4. dinosaurs: The children were pretending to be <dinosaurs>.
 - a. robbers who work at sea.
 - b. very small creatures with human form but with wings
 - c. large creatures with wings that breathe fire
 - d. animals that lived an extremely long time ago
- 5. scrub: He is <scrubbing it>.
 - a. cutting shallow lines into it.
 - b. repairing it
 - c. rubbing it hard to clean it
 - d. drawing simple pictures of it
- 6. cube: I need one more <cube>.
 - a. sharp thing used for joining things
 - b. solid square block
 - c. tall cup with no saucer

- d. piece of stiff paper folded in half
- 7. orientate: He needs to be <orientated>.
 - a. taught eastern culture
 - b. given training to improve his skills
 - c. made to do his work
 - d. made familiar with where things are
- 8. loo: He's in <the loo>.
 - a. serious difficulties
 - b. the toilet
 - c. the prison
 - d. the small building for tools in a garden
- 9. photocopy: The material was <photocopied>.
 - a. decorated with a picture made by a camera
 - b. recorded on film for use in a law court
 - c. copied by printing a picture of it made with a camera
 - d. copied many times
- 10. whoop: The audience <whooped>.
 - a. made happy loud cries
 - b. showed disappointment
 - c. clapped loudly
 - d. slowly got smaller

Thousand 4

- 1. devastate: The city was <devastated>.
 - a. made beautiful for a special occasion
 - b. cut off from the rest of the world
 - c. turned into ruins
 - d. made dirty and unpleasant by small animals
- 2. evacuate: They were <evacuated>.
 - a. moved to another place for safety
 - b. searched in case they had guns or knives
 - c. extremely frightened
 - d. made to look guilty of a crime
- 3. tummy: Look at my <tummy>.
 - a. cloth to cover the head
 - b. stomach
 - c. small furry animal
 - d. thumb
- 4. candid: Please <be candid>.
 - a. be careful
 - b. show sympathy
 - c. show fairness to both sides
 - d. say what you really think
- 5. mash: We <mashed> the food.
 - a. cooked
 - b. ate
 - c. crushed

- d. threw out
- 6. parallel: He drew two <parallel lines>.
 - a. sloping lines
 - b. lines the same distance apart everywhere
 - c. lines of equal length
 - d. lines that make a cross
- 7. marble: It was made of <marble>.
 - a. a kind of hard rock
 - b. a kind of hard wood
 - c. a kind of soft metal
 - d. woven pieces of string
- 8. weep: He <wept>.
 - a. finished his course
 - b. cried
 - c. died
 - d. worried
- 9. exert: Don't <exert yourself>!
 - a. praise yourself too much
 - b. hurt yourself
 - c. make yourself work hart
 - d. give yourself everything you want
- 10. frog: Look at the <frog>.
 - a. small tree with poisonous fruit
 - b. steam engine
 - c. animal that lives on land and in water
 - d. mass of low clouds

Thousand 5

- 1. whey: Can you use the <whey>?
 - a. watery part of milk
 - b. yellow part of an egg
 - c. very small pieces produced by cutting wood
 - d. skin of some fruits
- 2. legend: It is now <a legend>.
 - a. a building for keeping old things
 - b. a thing that is always done
 - c. an old story
 - d. a regular event
- 3. threshold: They raised the <threshold>.
 - a. flag
 - b. point or line where something changes
 - c. roof inside a building
 - d. cost of borrowing money
- 4. sheriff: The <sheriff> was friendly.
 - a. person who flies airplanes
 - b. person who takes care of babies
 - c. person who makes sure that the law is obeyed

- d. person who teaches children at home
- 5. premature: He was <premature>.
 - a. born earlier than expected
 - b. able to guess what would happen in the future
 - c. likely to decide about things without knowing much about them
 - d. able to do things younger than usual
- 6. warfare: Modern <warfare> is frightening.
 - a. crime
 - b. dancing
 - c. fighting
 - d. spoiling of the world
- 7. lavatory: Where is the <lavatory>?
 - a. place where books are kept
 - b. place where clothes are washed
 - c. toilet
 - d. place for science experiments
- 8. wilderness: It is <a wilderness>.
 - a. an exciting event
 - b. an event that is not easily explained
 - c. a place in its natural state
 - d. a place for throwing rubbish away
- 9. abundance: We have <an abundance>.
 - a. a problem that makes progress difficult
 - b. something that needs attention at once
 - c. an agreement to work together
 - d. a bigger quantity than we need
- 10. therapeutic: Doing this is <therapeutic>.
 - a. a waste of money
 - b. very wicked
 - c. helpful in curing illness
 - d. likely to make you go to sleep

Thousand 6

- 1. kink: There is a <kink> in it.
 - a. small hole
 - b. bend
 - c. hidden message
 - d. very small biting insect
- 2. cavalier: He treated her <in a cavalier manner>.
 - a. without care
 - b. politely
 - c. awkwardly
 - d. as a brother would
- 3. psycho: He was a <psycho>.
 - a. person who performs strange tricks with his body
 - b. person who claims to be able to speak with dead people
 - c. person whose mind is severely disturbed

- d. doctor who heals people by telling them things while they are asleep
4. shrug: She <shrugged>.
- cheated
 - moved her shoulders up and down
 - stayed longer than other people
 - lay on top of the water
5. absurd: This is <absurd>.
- unusual
 - not well known
 - hard to understand
 - unreasonable
6. genre: They argued about <its genre>.
- the class it belongs to
 - its origin or roots
 - the information it contains
 - whether it applies to everybody
7. bloc: They have joined this <bloc>.
- musical group
 - band of thieves
 - small group of soldiers who are sent ahead of others
 - group of countries with a common purpose
8. lucrative: It is <a lucrative business>.
- a business which makes a lot of money
 - a business with very honest habits
 - a business which breaks the law
 - a business which is not yet fully developed
9. peril: It is a time of <peril>.
- happiness
 - danger
 - hope
 - change
10. purport: He <purports to be> a doctor.
- wants to be
 - refuses to be
 - boasts that he is
 - claims that he is

Thousand 7

1. spook: He is afraid of <spooks>.
- ghosts
 - sea creatures that walk sideways
 - little bits of hot stuff which fly out of a fire
 - small snakes which hide in grass
2. tangible: The rewards were <tangible>.
- not worth having
 - very great
 - pleasing

- d. real
- 3. turbulent: The river was <turbulent>.
 - a. flowing quickly
 - b. unpleasant because it was unable to flow
 - c. full to the top of its banks
 - d. violently disturbed
- 4. erratic: He was <erratic>.
 - a. without fault
 - b. very bad
 - c. very polite
 - d. unsteady
- 5. null: His influence <was null>.
 - a. had good results
 - b. was unhelpful
 - c. had no effect
 - d. was long-lasting
- 6. beck: She sat by the <beck>.
 - a. small river
 - b. fire outside on the ground
 - c. edge of the road
 - d. round window
- 7. ravage: The forest was <ravaged>.
 - a. mapped carefully
 - b. badly damaged
 - c. protected
 - d. increased in size
- 8. furor: This caused <a furor>.
 - a. wild excitement
 - b. a large, very hot fire
 - c. a time when there is little food
 - d. the complete failing of plans
- 9. speckled: It is <speckled>.
 - a. covered in small spots
 - b. decorated with shiny metal circles
 - c. covered in rough lumps
 - d. like a saw, with teeth at the edge
- 10. rampage: We were surprised at the <rampage>.
 - a. escape of liquids or gases
 - b. distance traveled
 - c. wild or violent rushing about
 - d. amount of a medicine to be taken

A Controlled-production Vocabulary Size Test

(From Laufer and Nation, 1999)

Complete the underlined words. The example has been done for you.

He was riding a bicycle.

The 2000-word level

1. I'm glad we had this opp_____ to talk.
2. There are a doz_____ eggs in the basket.
3. Every working person must pay income t_____ .
4. The pirates buried the trea_____ on a desert island.
5. Her beauty and cha_____ had a powerful effect on men.
6. La_____ of rain led to a shortage of water in the city.
7. He takes cr_____ and sugar in his coffee.
8. The rich man died and left all his we_____ to his son.
9. Pup_____ must hand in their papers by the end of the week.
10. This sweater is too tight. It needs to be stret_____ .
11. Ann intro_____ her boyfriend to her mother.
12. Teenagers often adm_____ and worship pop singers.
13. If you blow up that balloon any more it will bur_____ .
14. In order to be accepted into the university, he had to impr_____ his grades.
15. The telegram was deli_____ two hours after it had been sent.
16. The differences were so sl_____ that they went unnoticed.
17. The dress you're wearing is lov_____ .
18. He wasn't very popu_____ when he was a teenager, but he has many friends now.

The 3000-word level

1. He has a successful car_____ as a lawyer.
2. The thieves threw ac_____ in his face and made him blind.
3. To improve the country's economy, the government decided on economic ref_____ .
4. She wore a beautiful green go_____ to the ball.
5. The government tried to protect the country's industry by reducing the imp_____ of cheap goods.
6. The children's games were funny at first, but finally got on the parents' ner_____ .
7. The lawyer gave some wise coun_____ to his client.
8. Many people in England mow the la_____ of their houses on Sunday morning.
9. The farmer sells the eggs that his he_____ lays.
10. Sudden noises at night sca_____ me a lot.
11. France was proc_____ a republic in the 18th century.
12. Many people are inj_____ in road accidents every year.
13. Suddenly he was thru_____ into the dark room.
14. He perc_____ a light at the end of the tunnel.
15. Children are not independent. They are att_____ to their parents.
16. She showed off her sle_____ figure in a long narrow dress.
17. She has been changing partners often because she cannot have a sta_____ relationship with one person.
18. You must wear a bathing suit on a public beach. You're not allowed to be na_____ .

The 5000-word level

1. Soldiers usually swear an oa_____ of loyalty to their country.
2. The voter placed the ball_____ in the box.
3. They keep their valuables in a vau_____ at the bank.
4. A bird perched at the window led_____ .
5. The kitten is playing with a ball of ya_____ .

6. The thieves have forced an ent_____ into the building.
7. The small hill was really a burial mou_____ .
8. We decided to celebrate New Year's E_____ together.
9. The soldier was asked to choose between infantry and cav_____ .
10. This is a complex problem which is difficult to compr_____ .
11. The angry crowd sho_____ the prisoner as he was leaving the court.
12. Don't pay attention to this rude remark. Just ign_____ it.
13. The management held a secret meeting. The issues discussed were not disc_____ to the workers.
14. We could hear the sergeant bel_____ commands to the troops.
15. The boss got angry with the secretary and it took a lot of tact to soo_____ him.
16. We do not have adeq_____ information to make a decision.
17. She is not a child, but a mat_____ woman. She can make her own decisions.
18. The prisoner was put in sol_____ confinement.

The University Word List level

1. There has been a recent tr_____ among prosperous families towards a smaller number of children.
2. The ar_____ of his office is 25 square meters.
3. Phil_____ examines the meaning of life.
4. According to the communist doc_____ , workers should rule the world.
5. Spending many years together deepened their inti_____ .
6. He usually read the sport sec_____ of the newspaper first.
7. Because of the doctors' strike the cli_____ is closed today.
8. There are several misprints on each page of this te_____ .
9. The suspect had both opportunity and mot_____ to commit the murder.
10. They insp_____ all products before sending them out to stores.
11. A considerable amount of evidence was accum_____ during the investigation.
12. The victim's shirt was satu_____ with blood.
13. He is irresponsible. You cannot re_____ on him for help.
14. It's impossible to eva_____ these results without knowing about the research methods that were used.
15. He finally att_____ a position of power in the company.
16. The story tells us about a crime and subs_____ punishment.
17. In a hom_____ class all students are of a similar proficiency.
18. The urge to survive is inh_____ in all creatures.

APPENDIX E

MATERIAL FOR GRAMMATICALITY JUDGMENT TASK

(adapted from Ellis, 2005)

1. Since (G) I haven't seen him for a long time.
2. Comparative (G) I think that he is nicer and more intelligent than all the other students.
3. Dative (G) The teacher explained the problem to the students.
4. V comp (UG) Liao says he wants buying a car next week.
5. past ed (UG) Martin completed his assignment and print it out.
6. Tag (UG) We will leave tomorrow, isn't it?
7. Adverb (G) He plays soccer very well.
8. Aux do (UG) Did Keiko completed her homework?
9. Modal (UG) I must to brush my teeth now.
10. Conditional (UG) If he had been richer, she will marry him.
11. Since (UG) He has been living in New Zealand since three years.
12. Reported (G) Pam wanted to know what I had told John.
13. Article (UG) They had the very good time at the party.
14. Passive (UG) Between 1990 and 2000 the population of New Zealand was increased.
15. Possessive (UG) Liao is still living in his rich uncle house.
16. Plural (UG) Martin sold a few old coins and stamp to a shop.
17. Since (UG) I have been studying English since a long time.
18. Modal (UG) I can to speak French very well.
19. Past ed (UG) Joseph miss an interesting party last weekend.
20. 3rd person s (G) Keiko eats a lot of sushi.
21. Reported (G) Bill wanted to know where I had been.
22. Aux do (G) Did Cathy cook dinner last night?
23. Dative (G) Rosemary reported the crime to the police.
24. Comparative (G) Mary is taller than her sisters
25. 3rd person s (UG) Hiroshi live with his friend Koji.
26. V comp (G) Keum wants to buy a computer this weekend.
27. Adverb (UG) She writes very well English.
28. Conditional (G) If she had worked hard, she would have passed the exam.
29. Reported (UG) Tom wanted to know whether was I going.
30. Article (UG) I saw very funny movie last night.
31. Dative (UG) The teacher explained John the answer.
32. Modal (G) I must finish my homework tonight.

33. Possessive (UG) Keum went to the school to speak to her children teacher.
34. Since (G) Keiko has been studying in Auckland for three years.
35. Comparative (UG) This building is more bigger than your house.
36. Tag (G) That book isn't very interesting, is it?
37. Passive (G) Her English vocabulary increased a lot last year.
38. Past ed (G) Hiroshi received a letter from his father yesterday.
39. Aux do (G) Does Keum live in Auckland?
40. Plural (G) Liao left some pens and pencils at school.
41. Conditional (UG) If he hadn't come to New Zealand, he will stay in Japan.
42. Comparative (UG) My car is more faster and more powerful than your car.
43. Possessive (G) Joseph flew to Washington to meet the President's advisor.
44. V comp (UG) Joseph wants finding a new job next month.
45. 3rd person s (G) Liao works very hard but earns very little.
46. Article (G) Japan is a very interesting country.
47. Modal (G) I can cook Chinese food very well.
48. Adverb (G) They enjoyed the party very much.
49. Tag (UG) The boys went to bed late last night, is it?
50. Reported (UG) She wanted to know why had he studied German.
51. Dative (UG) He reported his father the bad news.
52. Possessive (G) Keiko spoke to the professor's secretary.
53. Past ed (G) Liao stayed at home all day and finished the book.
54. Plural (G) Hiroshi found some keys on the ground.
55. Article (G) They did not come at the right time.
56. Conditional (G) If he had bought a ticket, he might have won the prize.
57. V comp (G) Martin says he wants to get married next year.
58. Passive (UG) An accident was happened on the motorway.
59. 3rd person s (UG) Keum lives in Hamilton but work in Auckland.
60. Adverb (UG) She likes always watching television.
61. Aux do (UG) Did Martin visited his father yesterday?
62. Passive (G) Something bad happened last weekend.
63. Plural (UG) Keum bought two present for her children.
64. Tag (G) She is working very hard, isn't she?
65. Relative (UG) The bird that my brother caught it has died.
66. Relative (UG) The boat that my father bought it has sunk.
67. Relative (G) The book that Mary wrote won the prize.
68. Relative (G) The car that Bill has rented is a Toyota.

APPENDIX F
ESSAY RATING SCALES

For English Essays:

Band	Score	Content & Organization	Language use
6 excellent	14- 15	Issues dealt with fully, Clear position, Substantive arguments, balanced ideas with full support and logical connection, strong control of organization	Excellent control of language with effective choice of words, sophisticated range of grammatical structures and vocabulary, few or no errors
5 very good	11- 13	Issues dealt with well, Clear position, Substantive arguments, generally balanced ideas with support and logical connection, good control of organization, occasional repetition, redundancy, or a missing transition	Strong control of language, reads smoothly, sufficient range of grammatical structures and vocabulary with occasional minor errors
4 good	8-10	Issues discussed but could be better developed, positions could be clearer and supported with more substantive arguments, appropriate organization, with instances of redundancy, repetition, and inconsistency	Good control of language with adequate range of grammatical structures and vocabulary, may lack fluidity, some grammatical errors that do not interfere with understanding
3 fair	5-7	Issues discussed, but without substantive evidence, positions could be clearer and arguments could be more convincing, adequate organization, ideas are not always balanced	Acceptable control of language with major errors and limited choice of structures & vocabulary, grammatical errors that interfere with understanding
2 poor	2-4	Inadequate development of topic, little substance, ideas confused or disconnected, lacks logical sequencing	Weak control of language with frequent errors and limited range of structures & vocabulary
1 very poor	0-1	No substance, Not pertinent, no organization, no logic	Little control of language, dominated by errors

For Chinese Essays:

等级	分数	内容与结构	语言运用
6	90-100	圆满完成写作任务 观点明确，内容充实，论据充分，并进行了有效扩展；篇章组织条理清晰，有效地运用了衔接手段。	语言流畅，句型、词汇覆盖面广，错误极少；文体得当，格式正确。
5	80-89	很好地完成写作任务 观点清楚，内容充实，论据充分；整体篇章结构清晰，注意使用衔接手段	语言流畅，句型、词汇覆盖面较广，在运用较复杂句型时会出错；文体和格式正确。
4	70-79	较好地完成写作任务 表达了观点，但可以更明确；有论据论证，但可以更充分展开；整体篇章结构基本清晰，运用了一些衔接手段。	有一些明显的语言错误，但不影响对作文的理解，句型、词汇覆盖面尚可；文体和格式基本正确。
3	60-69	勉强地完成写作任务 所表达的某些内容要点不切题；论证不够充分；内容的组织安排不够清楚。	有一些明显的语言错误，有的已影响对作文的理解，句型、词汇覆盖面有限；文体和格式存在不一致或不正确的地方。
2	50-59	少量地完成写作任务 所表达的内容要点不切题；没有展开论证；内容的组织安排条理不太清楚，结构较混乱。	有较多基本的语言错误，影响对作文的理解；文体和格式不一致或不正确。
1	< 50	基本没有完成写作任务 没有论点，没有论据；内容的组织安排混乱，无任何条理。	大量的基本语言错误，严重影响对作文的理解；文体和格式不正确。

English translation of the rating scales for Chinese Essays:

Band	Score	Content & Organization	Language use
6 excellent	90-100	Task fulfilled fully Clear position, Substantive arguments, balanced ideas with full support, strong control of organization, effective use of cohesive devices	Good flow of language, sophisticated range of grammatical structures and vocabulary, few or no errors, appropriate style and format
5 very good	80-89	Task fulfilled very well Clear position, Substantive arguments, generally balanced ideas with adequate support, good control of organization, good use of cohesive devices	Generally good flow of language, sufficient range of grammatical structures and vocabulary with occasional minor errors, correct style and format
4 good	70-79	Task fulfilled well Positions expressed but could be clearer and supported with more substantive arguments, appropriate organization with some use of cohesive devices	Adequate range of grammatical structures and vocabulary, may lack fluidity, some grammatical errors that do not interfere with understanding, generally good style and format
3 fair	60-69	Task fulfilled fairly Positions could be clearer and more focused, arguments could be more convincing, ideas not always balanced	limited choice of structures & vocabulary, obvious errors that interfere with understanding, some inconsistency in style and format
2 poor	50-59	Task fulfilled minimally Content not closely related to the topic, inadequate development of argument, ideas confused or disconnected, structure disorderly	Very limited range of structures & vocabulary, many errors that interfere with understanding, inappropriate style and format
1 very poor	<50	Task not fulfilled at all No points made, No evidence, no organization, no logic	Too many errors that interfere seriously with understanding, incorrect style and format

APPENDIX H

APPROVED INFORMED CONSENT FORM

Georgia State University
 Department of Applied Linguistics & ESL
 Informed Consent
 (November 2009)

Title: Cognitive Factors Contributing to Adult EFL Learners' L2 Writing Performance in Essay Writing

Principal Investigator: Sara Weigle
 Student P. I.: Yanbin Lu

I. Purpose:

You are invited to take part in a research study. The purpose of the study is to investigate some of the factors that are involved in the writing ability of English learners. You are invited because you are a college student and a learner of English as a foreign language. A total of 150 participants will take part in this study. Participation will require about three hours of your time over three sessions in October through November, 2009.

II. Procedures:

If you decide to participate, you will take a series of tests in three sessions of an hour each. The tasks include: essay writing in Chinese and English (30 minutes each), questionnaire after each essay writing (10 minutes each), two vocabulary tests (20 minutes each), memory tests in Chinese and English (20 minutes each), and a grammar test (20 minutes).

Memory tests and grammar test will be done individually on a computer (computer test), while the other tests will be done in a group with paper and pen in a classroom outside your regular class hours (paper & pen test). A total of three hours is needed. You will be paid 100 RMB for your time and effort.

In all these tests, the only person you will interact with is the student researcher.

III. Risks:

In this study, you will not have any more risks than you would in a normal day of life.

IV. Benefits:

Participation in this study may benefit you personally. You will get a chance to practice taking different sections of English tests. You can also find out how well you did on the tests by sending a request to Yanbin Lu at ylu@gsu.edu (If you have any questions or concerns regarding the results of the tests, please feel free to contact Yanbin as well). Overall, we hope to gain information that will help us understand what factors are more important for language learners to write well in a second language.



V. Voluntary Participation and Withdrawal:

Participation in research is voluntary. You do not have to be in this study. If you decide to be in the study and change your mind, you have the right to drop out at any time. You may skip test items or stop participating at any time. However, if you skip items or quit, we would not be able to tell you how well you did on the tests.

VI. Confidentiality:

We will keep your records private to the extent allowed by law. We will use a research code rather than your name on study records. Only the student researcher and her advisor will have access to the information you provide. Information may also be shared with those who make sure the study is done correctly (GSU Institutional Review Board, the Office for Human Research Protection (OHRP) and any sponsor I may have). Test materials will be stored in locked cabinets and data entered into password- and firewall-protected computer. The key (code sheet) to identify the research participant will be stored separately from the data to protect privacy and will be destroyed when this protocol expires. Your name and other facts that might point to you will not appear when we present this study or publish its results. The findings will be summarized and reported in group form. You will not be identified personally.

VII. Contact Persons:

Contact Sara Weigle at sweigle@gsu.edu or Yanbin Lu at ylu@gsu.edu if you have questions about this study. If you have questions or concerns about your rights as a participant in this research study, you may contact Susan Vogtner in the Office of Research Integrity at +1-404-413-3513 or svogtner1@gsu.edu.

VIII. Copy of Consent Form to Subject:

We will give you a copy of this consent form to keep.

If you are willing to volunteer for this research, please sign below.

Participant

Date

Principal Investigator or Researcher Obtaining Consent

Date

