

Copyright Undertaking

This thesis is protected by copyright, with all rights reserved.

By reading and using the thesis, the reader understands and agrees to the following terms:

1. The reader will abide by the rules and legal ordinances governing copyright regarding the use of the thesis.
2. The reader will use the thesis for the purpose of research or private study only and not for distribution or further reproduction or any other purpose.
3. The reader agrees to indemnify and hold the University harmless from and against any loss, damage, cost, liability or expenses arising from copyright infringement or unauthorized usage.

IMPORTANT

If you have reasons to believe that any materials in this thesis are deemed not suitable to be distributed in this form, or a copyright owner having difficulty with the material being included in our database, please contact lbsys@polyu.edu.hk providing details. The Library will look into your claim and consider taking remedial action upon receipt of the written requests.

**INTEGRATING EXTENSIVE READING INTO CHINA'S
ENGLISH LANGUAGE CURRICULUM:
THE CASE OF A SHANGHAI SENIOR HIGH SCHOOL**

HE MU

Ph.D

The Hong Kong
Polytechnic University

2013

The Hong Kong Polytechnic University

Department of English

**Integrating Extensive Reading into China's English
Language Curriculum:
The Case of a Shanghai Senior High School**

HE Mu

A thesis submitted in partial fulfillment of the requirements for the
degree of Doctor of Philosophy

June 2012

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

_____ (Signed)

He Mu (Name of Student)

Abstract

The primary purpose of this research undertaking was to investigate the significance of extensive reading on the reading, listening and writing proficiency and linguistic knowledge (grammar and vocabulary) of senior high school students in a sample school in Shanghai, Mainland China.

This study used the Control-Group Post-test Only Design. Two tests were employed to establish three comparable groups for this study in terms of English language proficiency. The first test was the pre-term stratification test developed by the test bureau of the school district, which was used to stratify students with almost equal English language proficiency. The second test was the Edinburgh Project on Extensive Reading (EPER) Placement Test, which ensured that the students who entered this study possessed approximately equal ability to read in English. Ninety-nine students were stratified into three groups of thirty-three on the basis of overall English language proficiency, reading proficiency and gender. The groups were: the Complementary Reading (CR) group, the Free Reading (FR) group and the Control Group (CG). The CR group and the FR group formed the experimental population. Extensive reading was integrated into the English language curriculum of the CR group as a component complementary to the normal English language curriculum, whereas extensive reading was used as a medium of additional input for the FR group. Intensive (rather than extensive) reading was used in the Control Group (CG). The first final-term English test of the school district served as the mid-test to monitor the students' progress halfway through the study. The second final-term test of the school district administered at the end of the first academic year served as the post-test at the end of the study; it was used to evaluate the effect of extensive reading on senior high school English language teaching and learning. The same EPER Placement Test used earlier as a pre-test for reading proficiency was re-administered to the students a few days later after the post-test at the end of this study (i.e. at the end of the first academic year). One year after the program ended, the final term test administered at the end of the second academic year functioned as the delayed post-test to collect data on the long-term effect of extensive reading.

An Analysis of Variance (ANOVA) test and post hoc Fisher's LSD test were conducted to locate the statistical mean differences on the post-test and delayed post-test between the Control Group (CG) and the two experimental groups (CR and FR). The paired-sample t-test was used to test the statistical mean differences on the pre- and post-EPER Placement tests.

The following major findings emerged from the data analysis:

1. The results of the locally developed pre- and post-tests and the EPER Placement Test suggest that extensive reading as both a medium of additional voluntary input for the FR group and as a required component complementary to normal English language curriculum for the CR group result in improvement in the sample students' overall English language proficiency.
2. The post-test results show improvement in specific proficiency areas (i.e. listening, reading, and writing proficiency) and language knowledge components (i.e. grammar and vocabulary). The delayed post-test results show students' further improvement in writing proficiency one year after the program ended; this suggests that it takes more time for extensive reading to show positive effect on the productive skills.
3. The analysis of the reading records suggest that extensive reading significantly improves students' reading speed.
4. Despite its benefits, extensive reading has limitations. It is less effective in helping students acquire some types of vocabulary knowledge, such as low frequency words. It was also less effective in helping students to improve their writing accuracy.

To summarize, the present study examined the effects of extensive reading in a Chinese context, and has shown that benefits generated from an extensive reading program integrated into standard curriculum. The teaching situation of the present study was challenging, spanning only one third of the whole senior high school period, in an EFL environment where the students lacked experience, habit and linguistic foundation for reading in English. In this challenging environment, extensive reading has nevertheless shown positive results. This provides solid evidence in favor of extensive reading.

Publications arising from the thesis

Book chapter

He, M. (in submission). Extensive reading and students' academic achievement: A case study. In P. Brown, T. Muller, & S. Herder (Ed.), *Fluency in EFL: Learning, teaching and research*. Routledge (New York).

Journal article

He, M. (in press). Does extensive reading promote reading speed? *The Reading Matrix*.

Conference proceedings

He, M., & Green, C. (2012). Challenges in implementing extensive reading in Shanghai senior high schools. *Extensive Reading World Congress Proceedings, 1*, 34-36.

Alavi, S., He, M., Kuriniawan, L. L., & Rivadeniera, G. (2012). Approaches to implementing extensive reading under difficult circumstances. *Extensive Reading World Congress Proceedings, 1*, 15-18.

He, M., & Yap, F. H. (2012). Suggested ways of providing an immersive environment in an EFL context: An extensive reading program in a Shanghai senior high school. *Proceedings of the 2nd international conference on foreign teaching and applied linguistics*, 132-133.

Acknowledgements

I would like to thank the following people for their guidance, support and kindness. Without them, the thesis can never be completed.

My deepest gratitude goes to my two chief supervisors, who guided me throughout my PhD study. Dr. Christopher Green, my first chief supervisor, persistently and patiently supported me in continued academic and research development from the outset when I was a novice and throughout the research process until completion of data collection. Dr. Green accepted me when I was a new hand. He spent countless hours guiding me in research, helping me to revise my writing word by word, and answering numerous questions. He has forever been a selfless, devoted, amiable as well as a humorous supervisor. Dr. Yap Foong Ha, my second chief supervisor, guided me through the writing up of the thesis. Dr. Yap spent countless hours with me to weigh every word in this thesis. On countless nights Dr. Yap stayed up with me until mid-night in the office just for teaching me more. I cannot not remember how many times she had to run for the last train. To my two chief supervisors, I am grateful for all the things they have taught me, their consistent support, their insightful views and invaluable advice which inspire me a lot, as well as their loving care to their supervisee. Under their guidance, my PhD study is one of the most enjoyable events in my life. The time I spent with them is one of the most precious memories in my life.

I am also indebted to Professor David Qian, who helped me achieve my dream in PolyU which changed the rest of my life. Professor Qian assumed multiple-roles during my whole PhD study: a distinguished scholar giving me insightful and forward-looking academic advice and an avuncular elder giving inspirations and unfailing support to the younger generation. I can never thank you enough, Professor Qian!

I would like to thank my two overseas supervisors from the University of Edinburgh, Professor David Hill and Dr. Aileen Irvine. I can never forget the teachings of Professor Hill, whose knowledge in extensive reading is second to none, and the treat of afternoon teas accompanied by conversations full of

wisdom and humor in his garden. I am also grateful to Dr. Irvine, who gave me various academic suggestions and improved the research design of this extensive reading program. My memory of Edinburgh is forever bright and warm.

I wish to extend my heartfelt gratitude to the two external examiners of my thesis, Dr. Robert Waring and Dr. Teresa Xu. I have benefited greatly from the incisive comments, critical ideas, advice and encouragement from these two research experts. Without their help, the entire effort would not have been as complete and as finely tuned as it has come to be.

I want to thank a number of people who read part or the whole thesis, gave me very insightful and constructive comments and enlightened me in various aspects. These willing commentators included Professor William Grabe, Professor Stephen Krashen and Professor Winnie Cheng. Aside from their willingness to look at chapters in this thesis, I have learned immensely from each of them.

I must express my gratitude to the participants and collaborators in this study, the ninety-nine students in No. 3 High School attached to East China Normal University in Shanghai, my colleagues Mr. Feng Haizhong, Miss Jing Zhe, Miss Lin Hongyu, Ms. Wang Shan, Ms. Wu Bei and Ms. Zhu Hui, and all the other colleagues in the English teaching and research group, the academic officer Mr. Zhang Lianhui and our principal Mr. Gan Yaqing, for their contribution and cooperation. Without them, this extensive reading program would have been impossible.

I am grateful to the Hong Kong Government for awarding a scholarship in support of my doctoral study, to the Oxford University Press for giving me a special book offer in support of this extensive reading program, and to the Extensive Reading Foundation for offering me a scholarship to attend the first Extensive Reading World Congress.

I am also grateful to my dear teachers in the Department of English. Dr. Li Lan gave me extra support and care during my whole study period in Hong Kong. Dr. Li has been giving me numerous invaluable suggestions for both research and

personal life. Her home is always a sweet place for us non-local students to find delicious food and family warmth. Dr. Stephen Evans presented excellent workshops to help us improve English and pursue academic excellence. Not only students in our department but also students in the whole university benefit hugely from his hard work.

Before concluding, I would like to extend a particular gratitude to my husband, who has been encouraging and supporting me in my further study in Hong Kong. He not only had to endure a lonely life when I was not home, but also had to soothe my nerves and bear my short-temper when I encountered difficulties. His endless love is a driving force in the past ten years and in the future. How lucky I am to have him as my lifetime partner!

Needless to say, all the remaining inadequacies in the thesis are my own responsibility.

Table of Contents

| | |
|---|-------------|
| Abstract..... | i |
| Acknowledgements..... | iv |
| List of Tables..... | ix |
| List of Abbreviations..... | xiii |
| 1 Introduction | 1 |
| 1.1 Reading in education | 1 |
| 1.2 English language teaching in China from 1978 to now | 5 |
| 1.2.1 English language teaching in the restoration phase (1978—1984) | 5 |
| 1.2.2 English language teaching in the development phase (1985—1991)..... | 12 |
| 1.2.3 English language teaching in the adjustment phase (1992—2000)..... | 20 |
| 1.2.4 English language teaching in the new and current round of curriculum reform (2001—now) | 25 |
| 1.3 Overview of the problem..... | 29 |
| 1.4 Finding a solution to the problem..... | 32 |
| 1.5 Purpose of this study | 37 |
| 1.6 Significance of this study | 38 |
| 1.7 Research questions | 39 |
| 1.8 Hypotheses | 40 |
| 2 Literature Review..... | 41 |
| 2.1 Characteristics of extensive reading..... | 41 |
| 2.1.1 Definitions of extensive reading..... | 41 |
| 2.1.2 Agreed characteristics of extensive reading | 45 |
| 2.1.3 A less-discussed issue: reading quantity vs. reading quality | 51 |
| 2.2 The research on extensive reading | 52 |
| 2.3 Reading materials..... | 59 |
| 2.3.1 Extensive reading by means of graded readers | 59 |
| 2.3.2 Extensive reading through other print texts..... | 69 |
| 2.3.3 Extensive reading through the Internet and other electronic media | 74 |
| 2.4 Extensive reading schemes..... | 76 |
| 2.5 Examples of extensive reading schemes in Asia | 77 |
| 2.5.1 The Hong Kong Extensive Reading Scheme in English (HKERS) | 77 |
| 2.5.2 Reading and English Acquisition Program (REAP) in Singapore | 79 |
| 2.5.3 Extensive reading in Japan | 81 |
| 2.6 Conclusion: guidelines for a school-based extensive reading program in Shanghai..... | 83 |
| 3 Design and Implementation of a School-based Extensive Reading Program in Shanghai | 87 |
| 3.1 Overview of the program | 87 |
| 3.2 Sampling procedures | 88 |
| 3.3 Instruments..... | 92 |
| 3.3.1 The pre-test..... | 94 |
| 3.3.2 The mid-test | 96 |
| 3.3.3 The post-test | 99 |
| 3.3.4 The delayed post-test..... | 99 |
| 3.3.5 The EPER Placement Test | 100 |
| 3.4 Reading materials..... | 103 |
| 3.5 Control of extraneous variables..... | 105 |
| 3.6 Treatments | 109 |
| 3.6.1 The Free Reading Treatment..... | 109 |
| 3.6.2 The Complementary Reading Treatment..... | 110 |
| 3.6.3 English language teaching in the Control group | 113 |
| 3.6.4 Monitoring..... | 114 |
| 3.6.5 Summary | 115 |

| | |
|--|------------|
| 3.7 Challenges in implementation..... | 116 |
| 3.8 Time frame | 120 |
| 4 Findings and discussion | 121 |
| 4.1 Data analysis | 121 |
| 4.2 Basic information of students' extensive reading..... | 123 |
| 4.3 The test results..... | 128 |
| 4.3.1 The pre-, mid-, post- and delayed post-tests..... | 128 |
| 4.3.2 The EPER Placement Test B..... | 140 |
| 4.4 Extensive reading and language knowledge..... | 142 |
| 4.4.1 Extensive reading and vocabulary acquisition | 142 |
| 4.4.2 Extensive reading and grammar acquisition..... | 157 |
| 4.5 Extensive reading and receptive skills | 180 |
| 4.5.1 Extensive reading and reading comprehension | 180 |
| 4.5.2 Extensive reading and reading fluency..... | 189 |
| 4.5.3 Extensive reading and listening proficiency..... | 198 |
| 4.6 Extensive reading and productive skills..... | 214 |
| 4.6.1 Extensive reading and translation..... | 214 |
| 4.6.2 Extensive reading and writing proficiency | 220 |
| 4.7 Findings from the delayed post-test | 227 |
| 4.8 Findings from the reading records..... | 231 |
| 4.8.1 Reading speed | 231 |
| 4.8.2 Preference for different genres | 241 |
| 5 Conclusion and recommendations..... | 245 |
| 5.1 Summary | 245 |
| 5.2 Limitations of the study..... | 247 |
| 5.3 Recommendations and further areas of inquiry..... | 251 |
| Appendices | 257 |
| Appendix A The pre-test..... | 257 |
| Appendix B The mid-test | 262 |
| Appendix C The post-test..... | 269 |
| Appendix D Reading list..... | 276 |
| Appendix E Consent Form (English Version) | 280 |
| Appendix F Consent Form (Chinese Version) | 282 |
| Appendix G A preliminary discussion on graded readers..... | 283 |
| References | 292 |

List of Tables

Chapter 1

| | |
|---|----|
| Table 1. 1 The national English syllabuses and curriculum standards issued between 1978 and 2003 | 6 |
| Table 1. 2 Description of the 1978 syllabus | 7 |
| Table 1. 3 Description of the 1980 syllabus | 9 |
| Table 1. 4 Results of the 1982/1983 sample tests..... | 11 |
| Table 1. 5 Composition of NMETs from 1978 to 1984..... | 12 |
| Table 1. 6 Description of the 1986/1988/1990 syllabuses..... | 14 |
| Table 1. 7 Composition of NMET from 1985 to 1990 | 20 |
| Table 1. 8 Description of the 1993/1996/2000 syllabuses..... | 21 |
| Table 1. 9 Description of the curriculum standards for senior high school in 2001 NECS.... | 25 |
| Table 1. 10 Results of the 2007 NMET in Zhang's (2011) study | 28 |
| Table 1. 11 Results of IELTS in Zhang's (2011) study..... | 28 |
| Table 1. 12 Vocabulary size required in the syllabuses in the four phases | 28 |
| Table 1. 13 Passing rate of the reading comprehension section in 2009/2010 NMETS..... | 29 |
| Table 1. 14 Frequency of rare words in various sources | 31 |
| Table1. 15 Time allocated for the English curriculum in Shanghai primary and secondary schools..... | 34 |
| Table 1. 16 Amount of extracurricular input required by the Shanghai Curriculum Standards | 34 |
| Table1. 17 Composition of the language proficiency areas and language knowledge components in 2004/2005/2008 NMETs..... | 36 |
| Table 1. 18 Composition of the latest NMET(s) | 36 |

Chapter 2

| | |
|--|----|
| Table 2. 1 Frequency of rare words in various sources | 64 |
| Table 2. 2 Components of an established extensive reading program..... | 76 |
| Table 2. 3 Concerns in implementing and running an extensive reading program | 83 |
| Table 2. 4 Reading experience and perceived positive effects of an extensive reading program | 83 |

Chapter 3

| | |
|---|----|
| Table 3. 1 Composition of the Free Reading group and the Control Group..... | 89 |
| Table 3. 2 Comparison of means of NMET(s) | 91 |
| Table 3. 3 Reliability statistics of the pre-test | 96 |
| Table 3. 4 Reliability statistics of the mid-test | 98 |

| | |
|--|-----|
| Table 3. 5 Reliability statistics of the post-test..... | 99 |
| Table 3. 6 Matching EPER levels with vocabulary levels, student levels and exam levels . | 101 |
| Table 3. 7 Results of piloted EPER Placement Test B in 2009 | 102 |
| Table 3. 8 EPER levels of Senior One students in the sample school in 2009/10..... | 104 |
| Table 3. 9 Descriptive statistics and results of ANOVA from the pre-test | 106 |
| Table 3. 10 Descriptive statistics of the SSHMT scores of the three groups..... | 107 |
| Table 3. 11 Multiple comparisons of the SSHMT scores by Fisher's LSD Test | 107 |
| Table 3. 12 Genders of the participants in the three groups | 107 |
| Table 3. 13 Summary of the treatments in the three groups | 116 |
| Table 3. 14 Timeframe of this school-based extensive reading program in 2010/11 | 120 |

Chapter 4

| | |
|---|-----|
| Table 4. 1 Format of participants' reading record..... | 122 |
| Table 4. 2 Detailed information of participants' extensive reading | 123 |
| Table 4. 3 Frequency of genres read by the Free Reading group and Complementary Reading group | 127 |
| Table 4. 4 Quantification of participants' extensive reading..... | 127 |
| Table 4. 5 Reexamination of participants' reading amount | 128 |
| Table 4. 6 Descriptive statistics and results of ANOVA from the pre-test | 130 |
| Table 4. 7 Test of homogeneity of variances from the pre-test | 131 |
| Table 4. 8 The results of multiple comparisons by the Fisher's LSD test in the pre-test..... | 132 |
| Table 4. 10 Test of homogeneity of variances from the mid-test | 133 |
| Table 4. 11 The results of multiple comparisons by the Fisher's LSD test in the mid-test | 135 |
| Table 4. 12 Descriptive statistics and results of ANOVA from the post-test | 137 |
| Table 4. 13 Test of homogeneity of variances from the post-test | 137 |
| Table 4. 14 The results of multiple comparisons by the Fisher's LSD test in the post-test | 138 |
| Table 4. 15 Descriptive statistics and results of ANOVA from the delayed-test | 139 |
| Table 4. 16 Test of homogeneity of variances from the delayed post-test (缺)..... | 140 |
| Table 4. 17 Results of the pre- and post-EPER Placement Test B..... | 140 |
| Table 4. 18 Group * Progress Cross tabulation of the EPER Placement Test of the whole sample | 141 |
| Table 4. 19 Chi-Square Tests of the results of the pre- and post-EPER Placement Test | 142 |
| Table 4. 20 Descriptive statistics and results of ANOVA from the vocabulary sections in the mid- and post-tests | 148 |
| Table 4. 21 The results of multiple comparisons by the Fisher's LSD test in the vocabulary sections in the mid- and post-tests..... | 148 |
| Table 4. 22 Detailed analysis of the ten words tested in the vocabulary section in the post-test | 152 |

| | |
|---|-----|
| Table 4. 23 Descriptive statistics and results of ANOVA from the vocabulary section in the post-test | 154 |
| Table 4. 24 The results of multiple comparisons by the Fisher's LSD test in the vocabulary section in the post-test | 155 |
| Table 4. 25 Guidelines for teaching grammar in the national syllabuses for English language teaching | 158 |
| Table 4. 26 Descriptive statistics and results of ANOVA from the grammar sections in the pre-, mid- and post-tests | 164 |
| Table 4. 27 The results of multiple comparisons by the Fisher's LSD test in the grammar sections in the pre-, mid- and post-tests | 165 |
| Table 4. 28 The grammar syllabus in the Shanghai English Curriculum Standards for Primary and Junior Middle School (2004) | 168 |
| Table 4. 29 Descriptive statistics and results of ANOVA from the grammar section in the post-test | 173 |
| Table 4. 30 The results of multiple comparisons by the Fisher's LSD test in the grammar section in the post-test | 174 |
| Table 4. 31 Oxford Bookworm grammar syllabus and levels | 179 |
| Table 4. 32 Descriptive statistics and results of ANOVA from the cloze and reading comprehension sections in the pre-, mid- and post-tests | 181 |
| Table 4. 33 The results of multiple comparisons by the Fisher's LSD test in the cloze and reading comprehension sections in the post-test | 182 |
| Table 4. 34 Descriptive statistics and results of ANOVA from the cloze section in the post-test | 184 |
| Table 4. 35 The results of multiple comparisons by the Fisher's LSD test in the cloze section in the post-test | 185 |
| Table 4. 36 Descriptive statistics and results of ANOVA from the reading comprehension section in the post-test | 186 |
| Table 4. 37 The results of multiple comparisons by the Fisher's LSD test on Items 66, 67 and 68 in the reading comprehension section in the post-test | 187 |
| Table 4. 38 Descriptive statistics and results of ANOVA from the title-matching subsection in the post-test | 188 |
| Table 4. 39 Pre- and post-EPER levels of the Complementary Reading group | 193 |
| Table 4. 40 Chi-Square test of the results of the pre- and post-EPER Placement Test of the Complementary Reading Group | 193 |
| Table 4. 41 Pre and post-EPER levels of the Free Reading group | 194 |
| Table 4. 42 Chi-Square test of the results of pre- and post-EPER Placement Test of the Free Reading Group | 194 |
| Table 4. 43 Pre- and post- EPER levels of the Control Group | 194 |

| | |
|---|-----|
| Table 4. 44 Chi-Square test of the results of the pre- and post-EPER Placement Test of the Control Group | 195 |
| Table 4. 45 Descriptive statistics and results of ANOVA from the listening sections in the mid- and post-tests | 200 |
| Table 4. 46 The results of multiple comparisons by the Fisher's LSD test in the listening section in the post-test | 201 |
| Table 4. 47 Comparison between the Fujian Program and the present study | 203 |
| Table 4. 48 Descriptive statistics and results of ANOVA from the listening section in the post-test | 207 |
| Table 4. 49 The results of multiple comparisons by the Fisher's LSD test on Item 12 in the listening section of the post-test | 210 |
| Table 4. 50 Suggested test points and formats for listening comprehension in the 2001 syllabus..... | 212 |
| Table 4. 51 Descriptive statistics and results of ANOVA from the translation sections in the pre-, mid- and post-tests | 216 |
| Table 4. 52 The results of multiple comparisons by the Fisher's LSD test in the translation sections in the pre-, mid- and post-tests | 216 |
| Table 4. 53 Descriptive statistics and results of ANOVA from the translation section in the post-test | 218 |
| Table 4. 54 The results of multiple comparisons by the Fisher's LSD test in the translation section in the post-test | 219 |
| Table 4. 55 Descriptive statistics and results of ANOVA from the writing sections in the pre-, mid- and post-tests | 224 |
| Table 4. 56 The results of multiple comparisons by the Fisher's LSD test in the writing sections in the pre-, mid- and post-tests | 225 |
| Table 4. 57 Descriptive statistics and results of ANOVA from the delayed-test | 228 |
| Table 4. 58 The results of multiple comparisons by the Fisher's LSD test in the delayed post-test | 230 |
| Table 4. 59 Amount of extracurricular reading required in the National English Curriculum Standard..... | 234 |
| Table 4. 60 Reading speeds at different levels of the two experimental groups..... | 235 |
| Table 4. 61 Increase in reading speed between each two levels of the two experimental groups..... | 236 |
| Table 4. 62 The least well-received five titles for Level 1, 2, 3 and 4..... | 243 |

List of Abbreviations

| | |
|-------------------|--|
| ANOVA | One-Way Analysis of Variance |
| AWL | Academic Word List |
| BNC | British National Corpus |
| CBEI | Content-based English Instruction |
| CET | College English Test |
| CLT | Communicative Language Teaching |
| COCA | Corpus of Contemporary American English |
| CR | Complementary Reading |
| EFL | English as a foreign language |
| EPER | Edinburgh Project on Extensive Reading |
| ER | Extensive Reading |
| ERP | Extensive Reading Program |
| ESL | English as a second language |
| Fisher's LSD test | Fisher's Least Significant Difference test |
| FVR | Free Voluntary Reading |
| IELTS | International English Language Testing System |
| FR | Free Reading |
| GSL | General Service List |
| GTM | Grammar-Translation Method |
| HKERS | Hong Kong Extensive Reading Scheme |
| L1 | first language |
| L2 | second language |
| MC | Multiple Choices |
| MOE | Ministry of Education |
| NECS | National English Curriculum Standards |
| NMET | National Matriculation English Test |
| OBW | Oxford Bookworm |
| REAP | Reading and English Acquisition Program |
| SPSS | Statistical Package for the Social Science |
| SSR | Sustained Silent Reading |
| SSSHMT | Shanghai Senior High School Matriculation Test |

| | |
|--------|---|
| SSS ER | Start with Simple stories Extensive Reading |
| TOEFL | Test of English as a foreign language |
| USSR | Uninterrupted Sustained Silent Reading |
| WFB | Word Frequency Book |

1 Introduction

1.1 Reading in education

A common way to begin a discussion on reading is to provide a definition of its concept. However, it is never easy to present a straightforward definition of reading in a single sentence. Many attempts have been made to define reading in a single sentence, e.g., “Reading is the ability to draw meaning from the printed page and interpret this information appropriately” (in Grabe & Stoller, 2011, p. 3), and “Reading refers to understanding a message from a writer” (Wagner, et al., 2006, p. 1112). There are at least several reasons why these simple definitions are inadequate.

Reading undeniably and incontrovertibly involves two necessary elements: a reader (referred to as reader variables in Alderson, 2000, including schemata and background knowledge, formal schemata, knowledge of genre/text type, meta-linguistic knowledge and meta-cognition, content schemata, knowledge of subject matter/topic, knowledge of the world, cultural knowledge, reader skills and abilities, reader purposes in reading, reader motivation and interest, reader affect, and other reader characteristics; for the differences between individual readers, see Koda, 2005), and a text (referred to as text variables in Alderson, 2000, including text topic and content, text type and genre, literary and non-literary texts, text organization, traditional linguistic variables, text readability, typographical features, verbal and non-verbal information, and the medium of text presentation; for text modification, see Blau, 1982; Claridge, 2005; Davis, 1984; Honeyfield, 1977; Leow, 1993, 1997, 2009; Nation & Wang, 1999; Widdowson, 1976; see also Section 2.3 for discussion on graded readers).

From the readers’ perspective, numerous skills and knowledge foundation act in combination to create the overall reading comprehension abilities (see Koda, 2005 for the components of skills and knowledge foundation). Reading comprehension activates a series of processes: (1) word recognition (recognition of words in the text), (2) syntactic parsing (extracting grammatical information from a group of words, which is crucial for disambiguating polysemy in word recognition), and (3) semantic proposition formation, i.e., building clause-level meaning on the basis of word recognition and syntactic parsing. This set of

processes is commonly referred to as bottom-up processing (Alderson, 2000) or lower-level processes (Grabe, 2009; Grabe & Stoller, 2011). The lower-level processes are language-oriented, which require the readers' knowledge about the language in which they read. Based on the knowledge-oriented lower-level processes, another set of processes that contributes to comprehension is activated. After the clause-level meanings are built, they are integrated into a broader network in which the clause-level meanings are connected and the main idea of the text is formed. This process is referred to as Text Model of Comprehension (Grabe, 2009), a summary of the text itself. Spontaneously, readers will interpret the text in terms of their reading purposes and background knowledge, etc. This process is called the Situational Model of Reader Interpretation, in which a number of reader variables noted above, such as knowledge of the world and cultural knowledge, have an important role to play. The description of how the lower-level and higher-level processes are carried out does not imply that they are carried out independently at different times. It is in working memory that these two sets of processes are integrated and comprehension occurs (see Section 4.5.3 for discussion on working memory and its role in comprehension).

This brief description of how reading comprehension occurs shows that one disadvantage of the definitions of reading noted above is that these definitions fail to present the knowledge foundation relying on which the readers draw meaning from the text. This disadvantage gives rise to a further issue that these definitions fail to propose a possibility that the readers' abilities to make sense of a text varies according to their proficiency of the languages in which they are reading, which will challenge how "appropriate" the interpretations of the texts are; similarly, another disadvantage of these definitions is that they do not describe how the reading skills work together to interpret the information and whether the "appropriate" skills are employed. From the text's perspective, the above definitions do not reflect the linguistic features of the texts, such as its readability and the intended readership. This lack of information on both readers and texts makes impossible the investigation into the interaction between readers and texts from a cognitive view.

In addition to the inadequacy in describing readers, texts and the interaction

between them, i.e., the inadequacy from a cognitive view of reading, these definitions do not specify for which purpose the reader reads. Reading is not an isolated activity that takes place in some vacuum but it is undertaken for some purposes in a social context (Alderson, 2000). These definitions do not convey an increasing tendency to view reading as a socio-cultural practice (e.g., Barton, 1994), and hinder the follow-up discussion on many issues, for example, reading in a first language and second language (for issues on L1 and L2 reading, see Hudson, 2007; Koda, 2007; Koda & Reddy, 2008).

The critical review of the above definitions brings to light the complex nature of reading and reading ability. Reading is not a simple and linear model in which the readers photocopy the texts into their brains and then interprets the texts into their desired meanings in a vacuum environment. Numerous components, including knowledge foundations, reading skills, reading strategies, reading purposes, social contexts, etc, interact to make reading comprehension occur. The implication from the discussion on the definition of reading is that there is so much involved in reading that a general definition is in danger of being trivial or banal. In most contexts and research, it is thus necessary to specify what sort of reading is being considered, for what purpose and in what context.

Compared with the complex nature of reading and reading ability, the importance of reading is more simple and straightforward. In modern society, reading is the door to an individual's successful personal and social development (Holden, 2004). It allows people to gain knowledge about the subjects that should be learned at school, including language, history, science, mathematics and other subjects (Lyon, 1997, cited in Clark & Rumbold, 2006). The learners' ability to read, write and communicate is collectively referred to as "literacy", which is the core concern of academic achievement in education (Zipperer et al., 2002). Among the three abilities, studies in both L1 and L2 have shown that reading contributes to the development of writing proficiency (e.g., Caruso, 1994; Lee & Hsu, 2009; Saleem, 2010). As Krashen (1993) points out, "we acquire writing style, the special language of writing, by reading" (p.72). Concerning the ability to communicate, studies in reading in the first language (L1) have shown that there are strong relations between the amount of reading and verbal fluency. In

fact, reading is the core in all academic disciplines (White, 2004).

As the core of all academic disciplines, reading is found to have considerable influence on other language skills and knowledge. Studies in L1 reading have shown that there are strong relations between the amount of reading and spelling, vocabulary knowledge, reading skills, verbal fluency and cultural knowledge. For example, the participant in Goodman's (1982) study learned to read and spell before going to school. She was able to spell correctly 58% of the words in the word list for Grade 3, and was able to recognize 91% of the words that were correctly spelled in the word list. After formal school education commences, there is a powerful relationship between amount of reading and vocabulary learning. It is estimated that L1 children pick up 1,000 to 4,000 words through reading per year (e.g., Graves & Watts-Taffe, 2002; Stahl, 1999).

Reading also plays a crucial role in the life of L2 learners. It is not only an important literacy skill that must be acquired for function purposes, but it is also a necessary means by which learners develop linguistic competence in an L2, a competence that can sub-serve a wide range of language skills, including, but not limited to, listening, speaking and writing (Han, et al., 2009), and also a source of vocabulary acquisition (e.g., Pulido, 2009). Renandya et al. (1999) made investigation into a group of variables in order to locate which may contribute to the learning gain in English. Among the following variables, including reading amount, the level of difficulty of the reading material, the degree to which the reading material was interesting and comprehensible, the participants' perception of extensive reading as being useful and pleasant, only extracurricular reading done was found to be significantly correlated with learning gains.

In addition to language knowledge, skills and overall proficiency, reading ability also largely influences learners' academic performance at school, for both L1 and L2 readers. As shown in the American National Assessment of Education (Kutner et al., 2007), there is a substantial gap between the eighth graders from middle-class families and those from the low-income families in terms of reading skills. The gap in reading skills leads to different results on academic performance. 43% of the students from low-income families, i.e., students who

were likely to be less competent in terms of reading skills, failed to attain the minimal standard of the national education assessment, whereas much fewer students (19%) from non-poor families, i.e., students who were likely to be more competent in terms of reading skills, performed on this poorly. In addition to academic achievement, the influence of reading is extended to students' future lives. Individuals with reading disabilities not only have academic difficulties during their school years but also lower educational attainment and consequently lower income in adulthood (Savolainen et al., 2008).

1.2 English language teaching in China from 1978 to now

As Cambourne (2002) suggests, "What is learned cannot be separated from the context in which it is learned" (p. 26). Equally importantly, what we want to learn cannot be based on a context into which we have not gained a deep insight. The introduction and discussion in this section focus on English language teaching in mainland China from 1978 to now, with an emphasis on the role of reading in the English language syllabus and curriculum.

China's thirty years of English language curriculum reform since 1978, namely the Reform and Opening Up, can be divided into four phases: restoration (1978—1984), development (1985—1991), adjustment (1992-2000) and the new-round curriculum reform (2001 to now) (Liu, 2008b). Since the end of the ten-year Cultural Revolution, the English curriculum in basic education has been given a good opportunity of restoration. Between 1978 and 2003, a total of eleven national syllabuses and two curriculum standards were issued by the Ministry of Education (MOE) and State Education Commission (SEC) as enumerated in Table 1.1 below.

1.2.1 English language teaching in the restoration phase (1978—1984)

The syllabuses issued in the restoration phase included the 1978, 1980 and 1982 syllabuses. The 1978 syllabus, the first national English language teaching syllabus issued after the end of the ten-year Cultural Revolution, divided the ten-year primary and secondary education into two equal five-year phases, and English classes began from Primary Three (Guan & Meng, 2007). The weekly sessions of English classes were 4-4-4-5-4-4-4-4, 1080 sessions of 40 to 45

Table 1. 1 The national English syllabuses and curriculum standards issued between 1978 and 2003

| Phase | Year | Names of the syllabuses |
|-----------------------------|----------|--|
| Restoration | 1978 | English Syllabus for the Full-time Ten-year Primary and Secondary Schools (Trial edition) ¹ (MOE, 1978) |
| | 1980 | English Syllabus for the Full-time Ten-year Primary and Secondary Schools (Revised edition) ² (MOE, 1980) |
| | 1982 | English Syllabus for the Six-year Full-time Key Secondary Schools (a draft for discussion) ³ (MOE, 1982) |
| Development | 1986 | English Syllabus for the Full-time Secondary Schools ⁴ (SEC, 1986) |
| | 1988 | English Syllabus for the Full-time Junior High Schools (Preliminary draft) ⁵ (SEC, 1988) |
| | 1990 | English Syllabus for the Full-time Secondary Schools (Revised edition) ⁶ (SEC, 1990) |
| Adjustment | 1992 | English Syllabus for the Full-time Junior High Schools (Trial edition) ⁷ (SEC, 1992) |
| | 1993 | English Syllabus for the Full-time Senior High Schools (Preliminary draft) ⁸ (SEC, 1993) |
| | 1996 | English Syllabus for the Full-time Senior High Schools (Trial edition) ⁹ (SEC, 1996) |
| | 2000 (a) | English Syllabus for the Full-time Junior High Schools for the Nine-year Compulsory Education (Revised edition) ¹⁰ (MOE, 2000a) |
| | 2000 (b) | English syllabus for the full-time senior high schools (revised edition) ¹¹ (MOE, 2000b) |
| New-round curriculum reform | 2001 | The English Curriculum Standards for the Full-time Senior High Schools of Compulsory Education (trial edition) ¹² (MOE, 2001) |
| | 2003 | The English Curriculum Standards for Senior High Schools (A draft for discussion) ¹³ (MOE, 2003) |

minutes in eight years. The 1978 syllabus was praised for playing an important role in eliminating the erroneous political effect on English language teaching, stabilizing the teaching order and ensuring the quality of teaching (Liu, 2008b). As shown in Table 1.2, the 1978 syllabus made clear four pairs of relationship that need to be correctly handled in English language teaching: political and ideological education versus language teaching, theory versus practice, input versus output, and teachers versus students. In terms of the input versus output relationship, it was made clear by the 1978 syllabus that priority should be given

¹ 《全日制十年制中小学英语教学大纲（试行草案）》

² 《全日制十年制中小学英语教学大纲（试行草案）第二版》

³ 《全日制六年制重点中学英语教学大纲（征求意见稿）》

⁴ 《全日制中学英语教学大纲》

⁵ 《九年制义务教育全日制初级中学教学大纲（初审稿）》

⁶ 《全日制中学英语教学大纲（修订本）》

⁷ 《义务教育全日制小学、初级中学课程计划（试行）》

⁸ 《全日制高级中学英语教学大纲（初审稿）》

⁹ 《全日制普通高级中学英语教学大纲（供试验用）》

¹⁰ 《九年义务教育全日制初级中学英语教学大纲（试验修订版）》

¹¹ 《全日制普通高级中学英语教学大纲（试验修订版）》

¹² 《全日制义务教育普通高级中学英语课程标准（实验稿）》

¹³ 《高中英语课程标准（征求意见稿）》

to input. The development of students' English reading proficiency was prioritized in English language teaching in basic education, as introduced in Table 1.2 that "Lay emphasis on cultivating students' English reading and self-learning ability with a certain degree of listening, speaking, writing and translation ability".

Table 1. 2 Description of the 1978 syllabus

| Items | Description |
|-------------------------|---|
| Guiding principles | <ol style="list-style-type: none"> 1. Correctly handle the relationship between political and ideological education and language teaching.¹⁴ 2. Correctly handle the relationship between theory and practice in English language teaching.¹⁵ 3. Correctly handle the relationship between listening/speaking and reading/writing.¹⁶ 4. Correctly handle the relationship between teaching and learning.¹⁷ |
| Guidelines for teaching | Lay emphasis on cultivating students' English reading and self-learning ability with a certain degree of listening, speaking, writing and translation ability in order to prepare the students for their further study and use of English in their work or further study in higher education institutions. ¹⁸ |
| Expected achievement | <p>After eight years' study, the students are expected to:</p> <ol style="list-style-type: none"> 1. have a command of basic phonetics and grammar rules;¹⁹ 2. have a vocabulary of 2,800 words and a certain number of frequently used expressions and phrases;²⁰ 3. be able to read English materials of average level of difficulty on general topics with the assistance of dictionaries and reference books;²¹ 4. have a certain degree of listening, speaking, reading and writing ability.²² |

To guide the teachers on reading instruction, the 1978 syllabus provided the following guidelines for teaching reading.

1. Reading instruction in class is text-based, supplemented by a variety of exercises.²³
2. Reading instruction based on text is far from sufficient for the development of students' reading proficiency. Students should be encouraged to read a

¹⁴ 正确处理英语教学中政治思想教育和语言教学的关系。

¹⁵ 正确处理英语教学中理论与实践的关系。

¹⁶ 正确处理听说与读写的关系。

¹⁷ 正确处理教与学的关系。

¹⁸ 着重培养学生的阅读能力和自学英语的能力，及一定的听，说，写和译的能力。为其毕业后在三大革命运动中进一步学习与运用英语或进入高校学习打好基础。

¹⁹ 掌握基本的语音和语法。

²⁰ 掌握 2800 个单词和一定数量的习惯用语。

²¹ 能借助词典阅读与课文难易相当的一般读物。

²² 具有一定程度的听，说，读，写能力。

²³ 阅读教学以课文为主，以各种练习为辅。

large amount of extracurricular reading materials in English under the guidance of teachers.²⁴

3. Teachers should make the students aware that the primary purpose of reading is to understand what they read rather than pay excessive attention to specific wording or quote out of context.²⁵
4. The benefits of reading such as enrichment of vocabulary and language knowledge and enhancement of the ability to use English can only be attained by reading in quantity, which is a gradual and long-term process.²⁶
5. Reading instruction should focus on helping students understand what they read. Asking the students to write summaries of their reading in Chinese or English and to answer the reading-related questions are effective means to guide students' reading and check their comprehension.²⁷
6. To obtain English reading and self-learning ability, students are advised to read simple original or adapted English materials as much as possible. In addition to grammar and vocabulary, students are advised to pay attention to the specific expressions, logical thinking, and background knowledge in English. These should also be the main concerns in reading instruction. Teachers should guide students how to refer to or make use of grammar books, dictionaries and other reference books, by which the students obtain the ability to solve problems independently.²⁸

After English language teaching was back on track after the implementation of the 1978 syllabus, revisions were made and the 1978 syllabus was finalized as the 1980 syllabus (see Table 1.3 below). Unlike the 1978 syllabus which lay direct emphasis on students' English reading and self-learning ability, the 1980 syllabus encouraged an integrated training in listening, speaking, reading and writing. The position of the 1980 syllabus on the input versus output relationship became blurred. The direct emphasis on reading was weakened to an integrated

²⁴ 单纯通过课文教学是难以达到这个教学目的的，教师还必须指导学生在课堂之外阅读大量的英语读物。

²⁵ 教师要使学生明确，阅读的主要目的是理解读物的内容，而不仅仅是猎取某些词句。

²⁶ 阅读的益处，诸如扩充词汇和语言知识，增强使用英语的能力，只有通过大量的阅读才能实现，这是这个长期逐渐的过程。

²⁷ 阅读教学应以帮助学生正确理解内容为重点。要求学生用中文或英文写出内容提要，或者回答有关内容的问题，是指导阅读和检查理解是否正确的有效方法。

²⁸ 学生要真正获得阅读和咨询能力，就必须尽可能多得阅读浅显的原著或经过改写的原著。这类读物的难点往往不只是在于语法和词汇，而在于英语特有的表达方法，逻辑思维，习惯用语以及各种背景知识。因此，在教学中应着重解决这些方面的问题，并指导学生查阅语法书记，习惯用语词典及其他工具书和参考书，使学生逐渐获得独立解决疑难问题的能力。

training in listening, speaking, reading, writing and translation.

Table 1. 3 Description of the 1980 syllabus

| Items | Guiding principles |
|-------------------------|---|
| Guiding principles | <ol style="list-style-type: none"> 1. The pattern of English language learning of Chinese learners must be studied and summarized. Only in this way can the teaching efficiency be improved, and the students can learn English that is conform to standard in a gradual manner.²⁹ 2. English language teaching in primary and secondary school should attach important to cultivating students' actual ability to use English.³⁰ 3. A foreign language can be better learned by means of the integrated training in listening, speaking, reading, writing and translation.³¹ 4. The leading role of teachers in English language teaching should be particularly emphasized.³² |
| Guidelines for teaching | Train the students in listening, speaking, reading, writing and translation with affordable emphasis on reading and self-learning ability in order to prepare them for further study and use of English. ³³ |
| Expected achievement | The same as the 1978 syllabus. |

Despite the difference, the students' achievement under the two syllabuses was remarkable. In 1982 and 1983, the same sample test was administered by the Ministry of Education to freshmen from 37 universities in seven provinces and municipalities including Beijing and Shanghai. The 1982 test included 2,530 freshmen and the 1983 test covered 3,054 freshmen. As described in Huang and Feng (1984), the sample test was divided into four sections: reading comprehension, vocabulary, grammar and cloze. All of the four sections employed a multiple choice format. In the reading comprehension section, there were three texts of slightly different levels of difficulty. The first two texts were of common themes and the third text was about popular science. The word count of the three texts was around 700. Each of the three texts was followed by several comprehension questions, which amounted to 25 comprehension questions in all. The word count of the 25 comprehension questions was approximately 800. In all, the word count of the reading comprehension section was 1,500. The test takers were requested to finish this section within 50 minutes. If a test taker was

²⁹ 英语教学要注意研究总结中国学生学习英语的规律。只有这样，才能提高教学效率，使学生循序渐进地学到符合规范的英语。

³⁰ 中小学英语教学应该重视培养学生实际使用英语的能力。

³¹ 学习一种外语，必须通过听，说，读，写，译的综合训练，才能较好地掌握。

³² 教师的主导作用在英语教学中应该特别加以强调。

³³ 对学生进行听，说，读，写，译等各方面的基本训练，一般侧重培养其阅读能力和自学英语的能力，为其进一步学习与运用英语打好基础。

able to answer 17 questions out of the 25 correctly, his or her reading speed was estimated to 1500 words/50 minutes=30 words per minute.

In the vocabulary section, there were 150 items. For each item, a Chinese word and four possible English equivalents were given. The test takers were requested to match the Chinese words with the correct English equivalents. For example,
heart A.热 B. 心 C. 难 D. 听

The 150 items in the vocabulary section were divided into two groups. The 100 items in the first group were developed from the first 5,000 words in Thorndike's *The Teacher's Word List of 10,000 Words* by systematic random sampling, i.e., starting from a randomly selected word and selecting the word between a certain interval, for example, every 10th word in 50. These 100 items were intended to measure the vocabulary size of the test takers. The latter 50 items in the second group were developed from the 2,100 words in the word list attached to the English textbooks compiled by the Ministry of Education for secondary schools. These 50 items were designed to reflect how the test takers, i.e., secondary school graduates, mastered the words taught in secondary school.

The 100 items in the grammar section were intended to constitute a comprehensive sample of English grammar structures. Specifically speaking, the grammar structures taught in secondary schools accounted for the vast majority of the items in the grammar section, and part of the grammar structures taught in university were also tested. The ration between easy, moderate and difficult items was 2:5:3. In the fourth section, cloze test, 25 words were deleted from a text of 260 words. For each blank, the test takers were required to restore the words by choosing from the four options given. This text was an excerpt of a scientist's biography.

As evidenced by the test results presented in Table 1.4 below, there was a substantial improvement in terms of freshmen's English language proficiency in terms of increase rate.

Table 1. 4 Results of the 1982/1983 sample tests in Fu's study

| Year | Vocabulary | Grammar | Cloze | Reading | Total |
|-------------------|------------|---------|-------|---------|--------|
| 1982 | 55.33 | 48.61 | 8.48 | 44.54 | 157.46 |
| 1983 | 59.30 | 55.29 | 9.00 | 49.86 | 173.45 |
| Increase rate (%) | 7.17 | 13.74 | 6.13 | 11.94 | 10.15 |

(Fu, 1984, p.219)

As the leading role of teachers in English language teaching was established, the dominant teaching model was “cramming” abstract knowledge, using the Grammar-Translation Method. As its name suggests, the Grammar-Translation Method lays great emphasis on grammar and translation, and the essence is to build accuracy and in turn to cultivate learners’ ability to construct correct sentences (Griffiths & Parr, 2001). This explains why the biggest increase was in grammar. Another merit of the Grammar-Translation Method is that it integrates grammar rules and comprehension, which plays an active role in improving students’ comprehension ability (Zhang, 2003). This explains why the second biggest increase was in reading.

The primary objection to the Grammar-Translation Method is its overemphasis on grammar and translation at the expense of communicative competence (e.g., Brown, 1994; Zuo, 2002). The overemphasis on language rules lowers the chance and reduces time which students could have spent on reading large amount of reading materials and developing other language skills. Despite the weaknesses of the Grammar-Translation Method, however, the National Matriculation English Test (NMET), the highest-stakes English examination in mainland China, sent to teachers and students a signal that was different from that in the syllabuses. Although emphasized in the syllabuses, reading had a low profile in the NMET. The proportion that reading accounted for in the NMET showed a declining tendency, from 20% in 1978 to 10% in 1985. On the contrary, the proportion of grammar in the NMET was stable, an approximately 20% proportion from 1979 to 1985. Further, some other test formats, such as interchange between sentence patterns and completion of the passage with the correct forms of verbs, also examined students’ grammar knowledge. It was roughly estimated that grammar accounted for at least 45% of the whole exam. In Table 1.5 below not all the subsections of the NMET are included in that the

composition of NEMT varied every year. Some subsections, such as filling the blanks, interchange between sentence patterns, were too vague to categorize because they actually were multi-purpose test formats that examined a variety of language skills and knowledge. For our discussion only the clear-cut sections such as phonetics, spelling, grammar, vocabulary, cloze, reading and translations are included.

Table 1. 5 Composition of NMETs from 1978 to 1984

| Year | Phonetics (point) | Spelling (point) | Grammar (point) | Vocabulary (point) | Cloze (point) | Reading Comprehension (point) | Translation (point) |
|------|----------------------|---------------------|--------------------|-----------------------|------------------|-------------------------------------|------------------------|
| 1978 | | | 30 | | 10 | 20 | 30 |
| 1979 | | | 20 | | 10 | 15 | 35 |
| 1980 | | 10 | 20 | | | 18 | 20 |
| 1982 | 8 | 12 | 20 | | | 18 | |
| 1983 | 8 | 10 | 20 | | | 14 | |
| 1984 | 5 | | 20 | 10 | | 10 | |

Besides the criticism on the dominant teaching method and the low proportion that reading constituted in the NMET, the position on reading in the two consecutive syllabuses is also noticeable. The 1978 syllabus laid direct and obvious emphasis on reading while the 1980 syllabus intended to recommend equal emphasis on and an integrated training in listening, speaking, reading and writing. Whether the shift from the priority given to reading from recommendations for integrated training in listening, speaking, reading and writing meant a decline in the importance of reading in the national syllabus will be examined in the subsequent syllabuses.

1.2.2 English language teaching in the development phase (1985—1991)

In the second phase, namely the development phase (1985-1991), the time allocated to primary and secondary education was adjusted by the Ministry of Education. The pattern of five years for primary education and five years for secondary education was adjusted to six years for primary education and three years for junior middle school education. The nine years' education was made compulsory for all school-age children by the Chinese government in 1986. Senior high school education lasts three years, and was not compulsory for all students.

In this phase, three national English syllabuses were issued. They were the 1986, 1988 and 1990 syllabuses as shown in Table 1.6 below. As reflected in the three consecutive syllabuses, the clear emphasis on reading was replaced by the focus on students' communicative competence and an integrated training in listening, speaking, reading and writing, with priority given to specific skills at different stages. The shift was considered to be a progress from emphasis on mono-ability, i.e., reading, to multi- or integrated ability, i.e., listening, speaking, reading and writing. Concerning the development of specific skills at different stages, English language teaching in secondary schools was divided into three stages as introduced in Hang (1993, p.55). The initial stage was from Junior One and Two, the transition stage included Junior Three and Senior One, and Senior Two and Senior Three belonged to the advanced stage. At the initial stage, English language teaching should give priority to listening and speaking. The main task at this stage was to develop listening and speaking ability. The rationale for the task at the initial stage was that oral input was prioritized over written input in light of the law or nature of learning a language. Listening and speaking were the basic forms of language communication. Training in listening and speaking would cultivate students' ability to think in English and use English in real communication situations.

Thus from the perspective of linguistics, students' age and psychological characteristics and teaching methodology, it was logical and scientific to start English language teaching with listening and speaking from the initial stage. The main task of the transition stage was to present a balanced training on listening, speaking, reading and writing. As a result, it strengthened students' gains in listening and speaking from the previous stage, and it prepared the students for the advanced stage which laid emphasis on reading. Based on the training of the first two stages, the students would have already had a certain degree of listening, speaking, reading and writing ability and acquired a certain number of English words and grammar rules by the time they move to the advanced stage. Based on this, the subsequent training in reading and increase in students' reading amount would be logical and smooth (Yi, 2010).

Table 1. 6 Description of the 1986/1988/1990 syllabuses

| Syllabus | Guidelines for teaching |
|----------|--|
| 1986 | <ol style="list-style-type: none"> 1. Give detailed instruction of basic language knowledge. Lay emphasis on cultivating students' ability to use the language for communication purposes.³⁴ 2. Encourage a balanced training in all proficiency areas with priority given to specific skills at certain stages.³⁵ 3. English is recommended as the dominant medium of instruction. Use Chinese when necessary.³⁶ |
| 1988 | <ol style="list-style-type: none"> 1. Strengthen basic language knowledge and training, giving priority to students' communicative competence.³⁷ 2. Give a comprehensive training in listening, speaking, reading and writing, giving priority to specific skills at different stages.³⁸ 3. English is recommended as the dominant medium of instruction. Use Chinese when necessary.³⁹ 4. Improve the quality of in-class instruction. Promote extracurricular activities.⁴⁰ 5. Increase the amount of language materials and practice. Improve the quality of English language teaching.⁴¹ |
| 1990 | The same as the 1986 syllabus. |

English language teaching at this stage was deeply influenced by the Audio-Lingual Method. During World War II, the audio-lingual method (or the structural approach) was developed by American structural linguists (Fries, 1945; Rivers, 1981; Twaddell, 1980). This approach suggests that most problems experienced by foreign language learners concern the conflict between different structural systems, and thus the Audio-Lingual Method divides English language teaching into five phases: recognition, imitation, repetition, variation and selection (Hang, 2000, p.21).

As summarized in Hang (2000), the characteristics of the Audio-Lingual Method include: (1) Lay emphasis on speaking, giving priority to listening and speaking. Speaking is of top priority in teaching, because reading and writing are based on listening and speaking. (2) Sentence pattern is the core of English language teaching. Students are trained to imitate, repeat and memorize the sentence patterns until the patterns are internalized. (3) Restrict the use of native language and translation. Encourage teaching in the target language as much as possible. (4) Correct the mistakes made by the learners promptly and strictly in order to

³⁴ 精讲语言基础知识，着重培养学生运用语言进行交际的能力。

³⁵ 综合训练，阶段侧重。

³⁶ 尽量使用英语，适当使用母语。

³⁷ 加强基础知识和基本训练，着重培养为交际运用语言的能力。

³⁸ 听，说，读，写全面训练，不同阶段略有侧重。

³⁹ 尽量使用英语，适当使用母语。

⁴⁰ 提高课堂教学质量，积极开展课外活动。

⁴¹ 增加语言材料和语言实践的量，提高英语教学的质量。

instill a habit of using the target language accurately. (5) Employ a wide range of modern audio-visual means for teaching purposes, such as power point, recording, films and TV, in order to strengthen input in different ways.

However, the time devoted to listening and speaking was not as positive as it was claimed. For all the intensive classroom instruction which aimed at teaching English structures, these structures which were learned discretely were not able to be put into use in real communication by the students. The time devoted to structure instruction was unproductive, and it spoilt interest and enjoyment (Elley & Mangubhai, 1981). In addition, the emphasis on listening and speaking generated negative influence on other language skills and language knowledge components. In May, 1985, a survey on the quality of English language teaching in fifteen provinces and municipalities was administered by the Ministry of Education. This survey was the largest-scale subject survey that had ever been conducted since the People's Republic of China was founded, lasting nineteen months. It was administered to 58,050 students and 1,614 teachers in 105 key secondary schools and 35 average secondary schools. The results showed that the students' English language proficiency was not significantly improved compared with the first phase. Taking vocabulary as an example, the vocabulary size of senior high school graduates was about 1,600 words in 1984, and in 1986 the number was 1,800 (Liu, 2008a). Compared with the requirements of the 1978/1980 syllabuses that the students were expected to have a vocabulary of 2,800 words when they graduated from senior high school, the shortfall of 1,000 words indicated that the time devoted to structure instruction and model of instruction did not achieve the intended purpose.

In the schools of the South Pacific, the situation was not different. A study conducted in Nieu in the late 1970s, approximately contemporary with the development phase (1985-1991) in China, showed the weakness of the Audio-Lingual Method. In this study, De'Ath (2001) introduced 50 highly interesting illustrated story books to primary three pupils. He called this the Shared Book or Shared Reading Method. One year later, the pupils in the Shared Book group outperformed the pupils in the Audio-Lingual group on all three measures: oral sentence repetition, reading comprehension and word recognition.

Although laying emphasis on listening and speaking skills, the Audio-Lingual Method was not as effective as the Shared Reading Method which did not allocate much time and instruction resource to listening and speaking.

The second characteristic of the development phase was the booming of Chinese-English bilingual education, as in the 1986, 1988 and 1990 syllabuses English was recommended as the primary medium of instruction for all school subjects including English. This form of bilingual education is known as Content-based English Instruction (CBEI). On the surface, the most attractive feature of CBEI is to integrate English language teaching with the teaching and learning of other school subjects, which entails the concurrent study of language and subject matter (Hu & Alsagoff, 2010). Theoretically, a moral justification for CBEI is made on the basis of bilingualism. A number of important benefits of bilingualism have been suggested by research, such as the bilingual's more broadened and flexible perspective by accessing culture, knowledge and ideological forms in two languages and the greater meta-linguistic awareness than monolinguals (e.g., Baker, 2006; Bialystok, 2007, 2009; Bialystok et al., 2005; Carlisle, et al., 1999).

Despite the claimed benefits, however, the suggested benefits of bilingual education can only be achieved under supportive conditions (Swain & Johnson, 1997). As suggested by the studies concerning the practice of bilingual education in mainland China (e.g., Pi, 2004; Shen, 2004; Zhang, 2002), the implementation of Chinese-English bilingual education was constrained by an extensive range of contextual and educational factors. The most important constraint on CBEI was that there was no sociolinguistic environment for English immersion in Chinese mainland: the vast majority of the Chinese do not have the need to communicate in English for socio-cultural purposes (Zhang, 2002). This constraint created a dilemma for bilingual education in this context. On one hand, schools were advised to create an English-rich environment (Zhang, 2003) in order to compensate for the lack of a favorable sociolinguistic environment; on the other hand, there was seemingly little external motivation for the education system to provide such environment.

The second constraint on CBEI was a serious lack of teachers who were able to teach other school subjects in English (Shen, 2004). The Chinese teacher education system trained the teachers to be either subject teachers or English teachers instead of bilingual teachers (Zhang, 2002). Specifically, the subject teachers did not have the oral English proficiency to teach non-language subjects in English; similarly, the English teachers did not usually have the content knowledge for other school courses such as physics and chemistry (Pi, 2004). It can be argued that the replacement of local English teachers with expatriate teachers should be a possible solution to this problem if conditions permitted. However, native speakers of English were not a panacea to the teacher problem. It was reported by Rao (2010) that the native English-speaking teachers' classroom performance was not entirely satisfactory to Chinese EFL students because they were not sensitive enough to detect students' linguistic problems or familiar with the local culture and education. This study weakens the claim that the teachers' oral English proficiency is the key to the success of CBEI. For either native-speakers or non-native speakers of English, there are certain obstacles for them to overcome before they can carry out CBEI successfully.

It was also important to evaluate the effect of CBEI in various areas. Because a coherent curricular structure and curriculum standards were made available to guide CBEI, there was a lack of a coherent and standardized evaluation system of CBEI (Hu & Alsagoff, 2010). The common concerns in evaluating CBEI came from the definition of bilingual education by Stern (1972) that bilingual education is "schooling provided fully or partly in a second language with the object in view of making students proficient in the second language while, at the same time, maintaining and developing their proficiency in the first language and fully guaranteeing their educational development" (cited in Swain, 2000, p.199). Based on this definition, the concerns for the evaluation of CBEI should included: (1) whether the students' development in their first language, i.e., Chinese, will suffer if all or most of the school courses are taught in English, (2) whether the students' content area knowledge would suffer in those subject areas which they study in English, and (3) whether the program was successful in terms of improving students' English language proficiency.

Regarding the first concern of whether the students' Chinese would suffer if all or most of the school courses were taught in English, there was no formal government report about the students' declining Chinese proficiency, but there were non-governmental complaints on the society's overemphasis on English. Some scholars and researchers were alarmed by the aggressive expansion of English in the education system. They proposed that English be removed from the higher education institution matriculation examination system in order to avoid its potential threat to Chinese, and this presupposed the students' declining Chinese proficiency. Regarding the second concern of whether the students' content area knowledge would suffer in subjects taught in English, Jin and Zhuang (2002) reported a case in which the students who had been taught mathematics in English for around two months performed badly in assessment and as a result they had to be re-taught the same content in Chinese. On teachers' side, the teachers in Pi's (2004) study also complained that the students were not capable of bilingual education in mathematics because of their relatively low English language proficiency, which in turn compelled the teachers to simplify or even reduce the curricular content. These examples suggested that CBEI exerted negative influence on the students' content area knowledge in those subject areas which were taught in English.

Regarding the third concern of whether the program was successful in terms of English language proficiency, Gu and Dong (2005) appealed for great caution in the promotion of CBEI in mainland China in view of the unsuccessful English immersion education in Hong Kong. Before 1998, Cantonese was exclusively used as the medium of instruction in all primary schools. English was the dominant medium of instruction in approximately 90% of secondary schools, and used as the only official language in almost all the universities in Hong Kong (except the Chinese University of Hong Kong). Over the past decades, there have been complaints on the students' declining English proficiency. At the university level, there was a substantial gap in terms of English language proficiency between the university students in Hong Kong and their peers in America and Australia (e.g., Hyland, 1997; Littlewood & Liu, 1996; Stone, 1999). At the secondary school level, the students were found to be weak in terms of vocabulary (McNeil, 1994), grammar and reading (Johnson & Yau, 1996), and

writing (Yu & Aktinson, 1988a). According to Yu and Aktinson (1988b), using English as the medium of instruction exerted rather negative influence on students' academic achievements. Another consequence of English medium instruction in Hong Kong was that the Chinese proficiency of Hong Kong secondary school students was much lower than that of their peers in mainland and Taiwan (Shi, 1994). One of the most important implications that can be drawn from the unsuccessful English medium instruction in Hong Kong was that both teachers and students should reach a threshold level of English language proficiency and preferably the English language proficiency of the teachers and students should be approximately equal; otherwise, the benefits of English medium instruction cannot be fully attained (Cummins, 1991). Judging this from a higher level, the development and implementation of a language policy should be compatible with the cotemporary social context and educational system; otherwise, an unrealistic or overaggressive policy will lead to serious consequences (Gu & Dong, 2005).

To summarize, for the benefits of CBEI to be fully achieved, supportive conditions are needed: teachers who are capable of bilingual education, a real need for English as the dominant medium of instruction and a favorable sociolinguistic environment for English immersion. Currently, the optimal conditions needed for CBEI noted above are not available in mainland China. The only justification of CBEI is to expose the students to English as much as possible so that their English language proficiency can be improved.

Compared with the enthusiasm for communicative competence among the syllabus designers and CBEI in the society, the NMET adhered to its own style without giving any priority to listening, speaking or CBEI. As shown in Table 1.7 below, one of the obvious changes in the NMET in this phase was the big and increasing proportion of cloze in the NMET, from ten points in 1978, to twenty or more points since 1986. Given that cloze is an alternative form of reading task, this change implied the increasing importance of reading in the NMET in this phase. The misalignment between the syllabus and the NMET was notable, as the syllabus stressed that priority should be given to students' communicative competence, while the NMET increasingly prioritized and gave a clear advantage

to reading.

Table 1. 7 Composition of the NMET from 1985 to 1990

| Year | Phonetics (point) | Grammar (point) | Vocabulary (point) | Cloze (point) | Reading Comprehension (point) | Translation (point) |
|------|----------------------|--------------------|-----------------------|------------------|-------------------------------------|------------------------|
| 1985 | 8 | 22 | 25 | | 10 | 30 |
| 1986 | 5 | 17 | 8 | 20 | 30 | |
| 1987 | 5 | 15 | 10 | 20 | 30 | |
| 1988 | 5 | 15 | 10 | 20 | 30 | |
| 1989 | 5 | 30 | | 25 | 20 | 15 |
| 1990 | 5 | 15 | | 25 | 20 | 15 |

1.2.3 English language teaching in the adjustment phase (1992—2000)

In the third phase, namely the adjustment phase, five national English syllabuses were issued, two for junior middle school and three for senior high school. The 1993, 1996 and 2000 (b) syllabuses were included for discussion because they were designed specifically for senior high schools.

As shown in Table 1.8 below, the three syllabuses tended to achieve a balance between input and output, with an affordable emphasis on input. As specified in the 1993 syllabus,

In senior high school English language teaching, an integrated training in listening, speaking, reading and writing is encouraged. On the basis of students' enhancement in listening and speaking, emphasis on cultivating students' reading ability is suggested. Reading helps to enlarge students' vocabulary and enrich their language knowledge, which help them further understand the social and cultural background of the English-speaking countries. Listening is an important means of understanding and absorbing oral information. Listening and reading belong to the input domain. The students will have good speaking and writing proficiency only when they have received a sufficient amount of input. In teaching, teachers are encouraged to enhance the students' training in listening and reading, particularly in extensive reading. In fact, the emphasis on input is of particular importance to English language teaching in our country. In the first place, English is learned by the students as a foreign language instead of a second language. The context is not conducive to English language learning. The emphasis on input helps to compensate for the contextual constraint. In the second place, language learning will experience the process from accumulation to a qualitative leap. Without substantial input, a qualitative leap can hardly be achieved”⁴²

⁴² 在高中英语教学中，听，说，读，写要综合训练，在进一步提高学生的听说能力之同时，应侧重培养他们的阅

(Institute of Curriculum and Textbook, 1999, p. 403-405).

Table 1. 8 Description of the 1993/1996/2000 syllabuses

| Syllabus | Guidelines for teaching |
|---------------|--|
| 1993 Syllabus | <ol style="list-style-type: none"> 1. Consolidate and expand students' basic language knowledge, develop their ability in listening, speaking, reading and writing, cultivate their oral and written ability to use English for communicative purposes on the basis of the nine-year compulsory education.⁴³ 2. Give priority to reading ability and enable the students to gain self-learning ability.⁴⁴ 3. Prepare the students for further study and use of English; subject them to ideological, patriotic and socialism education.⁴⁵ 4. Enhance students' understanding of the native countries of the language that they learn.⁴⁶ 5. Develop students' intelligence and improve their ability to observe, notice, think, memorize and associate.⁴⁷ |
| 1996 Syllabus | Approximately the same as the 1993 syllabus. Add: Enhance students' understanding of foreign countries, especially English-speaking countries, to stimulate their learning interest. ⁴⁸ |
| 2000 Syllabus | <ol style="list-style-type: none"> 1. Consolidate and expand students' basic language knowledge, develop their listening, speaking, reading and writing skills, cultivate their ability to use English for communicative purposes on the basis of the nine-year compulsory education.⁴⁹ 2. Give priority to students' reading ability in order to subject them to ideological, patriotic and socialism education.⁵⁰ 3. Enhance students' understanding of foreign countries, especially English-speaking countries.⁵¹ 4. Develop students' intelligence in learning process, cultivate their innovative spirit and practical ability, and (help them) form effective learning strategies, which lays the foundation for their sustainable development in the future.⁵² |

Concurrent with the emphasis on a balance between input and output in the national syllabuses, the most noticeable initiative to reform English language teaching in China in the third phase was the effort to apply communicative language teaching in the Chinese context (e.g., Gu, 2009; Shi, 1997). As

读能力。阅读有助于学生扩大词汇量, 丰富语言知识, 进而了解英语国家的社会文化背景; 听是理解与吸收口头信息的重要手段。听和读都属于输入, 只有达到足够的输入量, 才能确保学生有较好的说与写的能力。因此, 在英语教学中教师要尽可能地加强学生的听读训练, 尤其要注意增加他们的泛读训练。事实上, 对语言输入的足够重视对我国的外语教学具有特殊意义。首先, 英语对我国学生来说是作为外语而不是第二语言来学习的, 学生的语言环境较差。因此, 对语言输入的重视有利于我们弥补这方面的不足。其次, 语言的学习必须经历由量的积累再到质的飞跃这样一个过程。因此, 没有足够的输入, 学生难以达到语言学习的质的飞跃。

⁴³ 在义务教育阶段教学的基础上, 巩固和扩大学生的基础知识, 发展其听, 说, 读, 写的基本技能, 培养其在口头上与书面上初步运用语言进行交际的能力。

⁴⁴ 侧重培养其阅读能力, 并使其获得一定的自学能力。

⁴⁵ 为他们将来继续学习与使用英语切实打好基础, 是他们受到思想品德, 爱国主义及社会主义方面的教育。

⁴⁶ 增进他们对所学语言国家的了解。

⁴⁷ 发展他们的智力, 提高他们观察, 主义, 思考, 记忆, 想象以及联想等方面的能力。

⁴⁸ 与 1993 年大纲基本相同。增加: 增进对外国特别是英语国家的了解, 激发学生的学习兴趣。

⁴⁹ 在义务教育初中阶段英语教学的基础上, 使学生巩固和扩大英语基础知识, 发展他们的听, 说, 读, 写基本技能, 提高他们初步运用英语进行交际的能力。

⁵⁰ 侧重提高他们的阅读能力, 使其在学习英语的过程中受到思想品德, 爱国主义及社会主义等方面的教育。

⁵¹ 增进其对外国文化, 尤其是对英语国家文化的了解。

⁵² 在学习中发展他们的智力, 培养他们的创新精神与实践能力, 形成有效的学习策略, 为他们今后的可持续发展奠定基础。

specified in Table 1.8, the 1993 syllabus calls for training in listening, speaking, reading and writing to enable students to use English for communicative purposes.

Communicative language teaching, however, failed to achieve its intended purpose. Similar to CBEI, communicative language teaching was constrained by a comprehensive range of contextual and educational factors. The first important hindering factor of communicative language teaching was inherent in communicative language teaching itself. Communicative language teaching breaks the systematic teaching of grammar, which is to the detriment of students' grammatical competence. Grammar is an essential aspect of communicative competence (Canale & Swain, 1980). Unless the students can be more skilled masters of the syntax of a language system, they could not really get to use the language communicatively. In fact, a lack of grammar knowledge would severely limit communicative competence (Wilkins, 1976).

Beside the first hindering factor inherent in communicative language teaching itself, one of the most important potential constraints on the application of communicative language teaching in the Chinese context was the traditional Chinese culture of learning. Hu (2002) argued that communicative language teaching was not compatible with the Chinese culture of learning in several important aspects. The traditional Chinese culture of learning played a vital role in setting the roles of teachers and students and dividing the responsibilities between teachers and students. Chinese culture is regarded as part of the Confucian-heritage cultures (Smith & Smith, 1999), influenced by which the Chinese students are expected to learn for their own sake instead of showing off and are required to show respect to the authority (Zhu, 2003). As a result, the Chinese students were not accustomed to expressing their opinions, especially in the presence of their teachers and classmates for fear that they should be laughed at if they made errors. They received knowledge in a passive manner and they rarely challenged the authority because it was considered good-mannered to be obedient. It has been confirmed by research that it was personality rather than language proficiency that was more likely to cause speaking anxiety. In Wang's (2010) investigation into the speaking anxiety of Chinese EFL learners,

personality factors were identified as the primary causes for students' speaking anxiety, which accounted for more variables than language proficiency.

Besides the constraints from students' side, another constraint on communicative language teaching came from the teachers' side: a severe shortage of qualified English teachers. Only 89.4% of the English teachers teaching junior middle school students and 55.0% of the teachers working in senior high school teachers in mainland China were considered professionally qualified (Yu, 2001). A considerable number of teachers only knew some basic English grammar and vocabulary. This partially explained why the Grammar-Translation Method was practiced so widely and survived so long, because teachers of low English language proficiency could basically teach English in Chinese using this method.

The teacher constraint did not only lie in teachers' English language proficiency but also other aspects. Besides qualification, the success of communicative language teaching largely depended how teachers talked and how teachers and students communicated with each other. In Liu and Zhao's (2010) study, the communication between teachers and students was directly influenced by how teachers talked. Among the choices in initiating an interaction, this type of question was employed most frequently by teachers but the least preferred by students. Direction was preferred by students but was not preferred by teachers. When producing no answer or an incorrect answer, students preferred to be informed by teachers but teachers usually prompted the students in class. When producing the correct answers, students looked forward to comments or more elaborate feedback but rather disappointedly teachers usually simply acknowledged the correct answers. In general, the teachers' language in the process of communication was not conducive to promoting more interaction or communication between teachers and students. These comparisons and the gap between the teachers and students indicated the gap between the ideal and reality of communicative language teaching in China.

As a result of the constraints noted above, the effect of communicative language teaching on English language teaching in mainland China was not satisfactory. In terms of speaking proficiency alone, a set of data from International English

Language Testing System (IELTS), one of the most reliable and authoritative English tests in the world, showed that the average score of the speaking section of the Chinese test takers was 5.36, which ranked the last in the world. In 2008 and 2009, the Chinese test takers still made no progress. In real communication situations, there were even more problems. The primarily intended purpose of communicative language teaching was to make up for the pragmatic functions, which meant being useful in real-world communication, ignored by the structure-based teaching approach, but it failed in this aspect. It just made students remember the isolated expressions, scenes and the correspondence between them. Furthermore, students had problems in using the limited expressions in real communication. Academic research and mass media reported that Chinese EFL learners adopted inappropriate communicative strategies. They asked ladies “How old are you?”, or asked strangers “What is your income?” etc. These were typical real life scenes that Chinese EFL learners would be involved in, but they seemed to be awkward in these scenes as shown by the above examples. More examples which showed the problems in real communication caused by the limitations of communicative language teaching can be found in research literature. For example, most Chinese undergraduates studying English were not English majors, so they needed to use communication strategies to facilitate their communication because they did not have sufficient exposure to the English language on a daily basis. Results from an investigation that analyzed communication strategies used by undergraduates non-English majors studying English in higher education institutions showed that students’ adoption of communication strategies was influenced by three variables: task type, English proficiency level and academic major (An, 2010). In Liu’s (2010) study, learners with higher English language proficiency level were more likely to adopt appropriate communication strategies. For example, in response to compliments, English majors were more likely to adopt *acceptance* strategies than non-English majors, whereas non-English majors were more inclined to adopt *non-acceptance* strategies than English majors. These studies suggested that communicative language teaching was not the single most effective way to enhance students’ communicative competence, but on the contrary the original English proficiency level was correlated with students’ communicative competence.

The most important pedagogical implication drawn from the implementation of communicative language teaching is that input is prior to output, although the syllabuses in the third phase tended to achieve a balance between input and output. For output to emerge, the students should receive a considerable amount of input and reach a threshold level, as shown in Wilkins (1976) that a lack of grammar knowledge would severely limit communicative competence and in An (2010) that English proficiency level was one of the most important variables influencing students' communication strategies.

1.2.4 English language teaching in the new and current round of curriculum reform (2001—now)

In the fourth phase, the new and current round of curriculum reform, the syllabus was changed into a set of curriculum standards. This change is considered a big progress and a milestone in Chinese curriculum reform by researchers. The obvious difference between a syllabus and a curriculum standard is that a curriculum standard prescribes clear requirements for students to attain without giving constraints on how to attain them (see Table 1.9 below). In other words, a curriculum standard allows more flexibility and room for material writing, classroom teaching and testing and assessment (Hu, 2005a).

Table 1. 9 Description of the curriculum standards for senior high school in 2001 NECS

| Level | Curriculum standards for each level |
|---------|--|
| Level 6 | <p>The students should:</p> <ol style="list-style-type: none"> 1. further enhance their motivation to learn English, and have a strong self-learning awareness;⁵³ 2. be able to understand the points of view in oral or written materials and state their own opinions;⁵⁴ 3. be able to describe personal experience in oral or written language effectively;⁵⁵ 4. be able to plan, arrange and carry out a variety of English language learning activities with the assistance of the teachers;⁵⁶ 5. be willing to expand and utilize learning resources and obtain information from various tunnels;⁵⁷ 6. be able to adjust learning objectives and strategies in light of self-evaluation;⁵⁸ 7. be able to realize and understand the cultural background of the language used in communication.⁵⁹ |
| Level 7 | <p>The students should:</p> |

⁵³ 进一步增强英语学习动机，有较强的自主学习意识。

⁵⁴ 能理解口头或书面材料中发表的观点并发表自己的见解。

⁵⁵ 能有效的使用口头或书面语言描述个人经历。

⁵⁶ 能在老师的帮助下计划，组织和实施各种英语学习活动。

⁵⁷ 能主动扩展和利用学习资源，从多渠道获取信息。

⁵⁸ 能根据自我评价结果调整学习目标和策略。

⁵⁹ 能体会交际中语言的文化内涵的背景。

| | |
|---------|--|
| | <ol style="list-style-type: none"> 1. have a clear and sustainable motivation to learn English and self-learning awareness;⁶⁰ 2. be able to communicate information on a wide range of topics, raise questions and state personal opinions and suggestions;⁶¹ 3. be able to read adapted readers for senior high school reading and English newspapers and magazines;⁶² 4. have basic practical writing skills;⁶³ 5. make use of the education resource for learning English;⁶⁴ 6. have a strong self-regulatory capacity. Form learning strategies that are suitable for individuals;⁶⁵ 7. understand the cultural differences in communication, and form basic cross-cultural awareness.⁶⁶ |
| Level 8 | <p>The students should:</p> <ol style="list-style-type: none"> 1. have relatively strong self-confidence and self-learning ability;⁶⁷ 2. be able to communicate smoothly with English speakers on familiar topics; be able to state evaluative views on oral or written materials;⁶⁸ 3. be able to write coherent essays in a complete format;⁶⁹ 4. be able to plan, arrange and carry out language practice activities independently, such as discuss and make plans, and report the findings of experiments and surveys;⁷⁰ 5. be able to make effective use of a variety of education resources such as the Internet to obtain and process information;⁷¹ 6. be able to evaluate learning outcomes on one's own initiative and form effective English learning strategies;⁷² 7. understand the cultural content and background in communication, and adopt a respect and all-embracing attitude towards cultures of different countries.⁷³ |
| Level 9 | <p>The students should:</p> <ol style="list-style-type: none"> 1. have self-learning ability;⁷⁴ 2. be able to listen and understand speeches, discussions, debates and reports on familiar topics;⁷⁵ 3. be able to exchange information and state personal attitudes and opinions on topics of common concern such as environmental protection, population, peace and development in English;⁷⁶ 4. be able to conduct oral translation in daily life;⁷⁷ 5. be able to conduct authentic communication in English on appropriate occasions;⁷⁸ 6. be able to read and understand popular science and literature on a wide range of topics with the assistance of dictionaries;⁷⁹ 7. be able to accomplish basic practical writing tasks, and obtain the initial ability to use the literature;⁸⁰ 8. be able to open up channels of self-learning, and enrich learning resources; have a strong sense of the world.⁸¹ |

⁶⁰ 有明确和持续的学习动机及自主学习意识。

⁶¹ 能就较广泛话题交流信息，提出问题并陈述自己的意见和建议。

⁶² 能读懂供高中学习阅读的英语原著改写本及英语报刊。

⁶³ 具有初步的实用写作能力。

⁶⁴ 利用教育资源进行学习。

⁶⁵ 具有较强的自我调控能力，初步形成适合自己的学习策略。

⁶⁶ 理解交际中的文化差异，初步形成跨文化交际意识。

⁶⁷ 有较强的自信心和自主学习能力。

⁶⁸ 能就熟悉的话题与讲英语的人士进行比较自然的交流，能就口头或书面材料的内容发表评价性见解。

⁶⁹ 能写出连贯且结构完整的短文。

⁷⁰ 能自主策划，组织和实施各种语言实践活动，如商讨和制订计划，报告试验和调查结果。

⁷¹ 能有效利用网络等多种教育资源获取和处理信息。

⁷² 能自觉评价学习效果，形成有效的英语学习策略。

⁷³ 了解交际中的文化内涵和背景，对异国文化采取尊重和包容的态度。

⁷⁴ 自主学习能力。

⁷⁵ 能听懂有关熟悉话题的演讲，讨论，辩论和报告的主要内容。

⁷⁶ 能就国内普遍关心的问题如环保，人口，和平与发展等用英语进行交谈，表明自己的态度和观点

⁷⁷ 能做到日常生活的口头翻译。

⁷⁸ 能利用各种机会用英语进行真实交际。

⁷⁹ 能借助字典阅读题材较为广泛的科普文章和文学作品。

⁸⁰ 能用常见应用文体完成一般的写作任务，并具有初步使用文献的能力。

⁸¹ 能自主开拓学习渠道，丰富学习资源，具有较强的世界意识。

Adopting the international grading standards, the 2001 National English Curriculum Standards (NECS) sets nine ability levels according to students' English language proficiency. Level 2, 5 and 8 are the graduation requirements from primary, junior high and senior high schools respectively. Level 3, 4, 6 and 7 are the transition levels between Level 2, 5 and 8.

Based on the 2001 NECS, the curriculum reform in senior high schools all over the country started. Zhang (2011) investigated the English language proficiency of two groups of students' to examine the effect of NECS after several years of its trial. The students in the first group were 120 freshmen randomly selected from thirteen departments in Fujian College of Engineering. The students in this group experienced the English curriculum before the new round of curriculum reform. The students in the second group were 150 freshmen randomly selected from the same college. The students in this group experienced the curriculum after the new round of curriculum reform. The two groups were identical in light of their total scores obtained in the university matriculation tests. The two assessment tools employed in this study were the 2007 NMET and the International English Language Testing System (IELTS). The reason for locating the 2007 NMET was that Fujian Province was authorized by the Ministry of Education to develop its own Matriculation English Test in 2004. Administering the 2007 national version of matriculation English test to freshmen from Fujian Province controlled and eliminated a number of extraneous variables such as local English syllabus and local English textbooks. IELTS is a standardized test of English language proficiency, whose reliability and validity has been internationally acknowledged. In IELTS, the test takers are required to take four sub-tests: listening, reading, writing and speaking. These four sub-tests are scored on a nine-band scale, with 9 the highest score. These four sections account for the same proportion in calculating the final score. To facilitate the research, Zhang set the full mark of each section at 40, the total mark being 160.

The two tests were administered to the 2007 group in 2008 and to the 2009 group in 2010. As shown in Table 1.10, the 2007 group outperformed the 2009 group in the NMET in terms of average score but the difference between them was not statistically significant ($F=0.058$, $p=0.810$). These results were reasonable given

that the 2009 Group experienced the new English curriculum which was considered to have been greatly improved.

Table 1. 10 Results of the 2007 NMET in Zhang's (2011) study

| Group | Average | F | Sig | Reading | F | Sig |
|-------|---------|-------|-------|---------|------|------|
| 2007 | 113.94 | 0.058 | 0.810 | 31.71 | 3.30 | 0.07 |
| 2009 | 113.55 | | | 32.73 | | |

* The mean difference is significant at the 0.05 level.

More interesting results emerged from the IELTS. As shown in Table 1.11 below, the 2007 group significantly outperformed the 2009 group in reading in IELTS, ($F=22.27$, $p<0.05$) although the total score of the 2009 group was significantly higher than that of the 2007 group ($F=11.54$, $p=0.001$).

Table 1. 11 Results of IELTS in Zhang's (2011) study

| Modules | 2007 | 2009 | F | Sig |
|-----------|-------|-------|-------|-------|
| Listening | 7.65 | 9.08 | 6.50 | <0.05 |
| Reading | 18.69 | 16.09 | 22.27 | <0.05 |
| Writing | 17.87 | 24.30 | 98.48 | <0.05 |
| Speaking | 24.59 | 25.04 | 0.93 | 0.336 |
| Total | 68.80 | 74.84 | 11.54 | 0.001 |

* The mean difference is significant at the 0.05 level.

Compared with the syllabuses issued in the first three phases, NECS sets higher requirement on students' vocabulary: students are required to acquire a vocabulary of 3,300 and a certain number of collocations and idioms as specified in the Level 9 requirement, while the requirement in the previous 2000 syllabus is 1950 (see Table 1.12 below). Given the strong correlation between vocabulary and reading comprehension (Qian, 1999, 2002; Qian & Schedl, 2004), it was reasonable to assume that the 2009 group could have done better in reading.

Table 1. 12 Vocabulary size required in the syllabuses in the four phases

| Phase | Syllabus | Junior High School | Senior High School |
|-------------------|----------|--------------------|--------------------|
| Restoration | 1978 | 1250 | 2200 |
| Development | 1986 | 1250 | 2750 |
| | 1992 | 1000 | / |
| Adjustment | 1993 | / | 3000 |
| | 2000 | 1200-1300 | 1950 |
| Curriculum reform | 2001 | 1500-1600 | 3300 |

(adapted from Huang, et al., 2004, p. 3)

The reason why a group with a smaller vocabulary size outperformed a group

with a bigger vocabulary size in IELTS reading was well worth discussing. Research suggests that more than one component contributes to the development of reading comprehension. These components include (1) syntactic knowledge and processing skills (e.g., Perfetti, et al., 2005; Stothard & Hulme, 1992; Tunmer & Hoover, 1992), (2) reading strategies that support comprehension (e.g., Schoot et al., 2008), (3) the complementary of reading strategies and higher-level processing to develop the strategic reader (e.g., Perfetti, et al., 2001; Pressley & Harris, 2006), (4) the role of discourse knowledge (e.g., Carrell, 1992; Williams, 2007) and the importance of vocabulary knowledge (e.g., Carver, 2003; Hazenbun & Hulstijn, 1996; Qian, 1999, 2002; Qian & Schedl, 2004; Thorndike, 1973) (see also Section 1.1 for the lower- and higher-level processes in reading). Important as it is, vocabulary knowledge is not the only deciding factor for reading ability. A larger vocabulary does not necessarily mean higher reading ability.

Zhang's study was not a unique case showing the decline of students' reading comprehension. The situation is the same in the National Matriculation English Test (Shanghai version) (NMETS). As shown in Table 1.13, the passing rate of all the reading subsections showed a declining tendency.

Table 1. 13 Passing rate of the reading comprehension section in the 2009/2010 NMETS

| Subsections | 2009 | 2010 |
|----------------|------|------|
| Cloze | 0.62 | 0.57 |
| Reading | 0.67 | 0.66 |
| Title Matching | 0.71 | 0.68 |
| Q&A | N/A | 0.61 |
| Total | 0.66 | 0.63 |

Note: The data in Table 1.12 were conveyed to teachers via the head of the English teaching staff in each school district from the examination committee.

1.3 Overview of the problem

As introduced and discussed in the previous section, the thirty years since the reform and opening up witnesses the change in China's English language curriculum and the consistent effort to present quality English education to the students. The effort made in the first phase was to cultivate students' reading ability, with the justification that reading helps the students understand and absorb written information, enlarge vocabulary, enrich language knowledge and then understand the cultural and social background of English-speaking countries.

As noted in Section 1.1, reading is the core in all academic disciplines (White, 2004), which is found to have a considerable influence on other language skills and knowledge. Studies in L1 reading have shown that there are strong relations between the amount of exposure to print input and cultural knowledge, vocabulary knowledge, reading skills, spelling, orthographic knowledge of words and verbal fluency, and studies in L2 reading reach the same conclusion. Under such a situation, reading should have gained prominence in English language curriculum, but the role of reading in syllabus and English language curriculum has been blurred.

The explanation for the unpopularity of reading in this phase is that reading demands a lot of resources, including class libraries and multiple copies of reading materials, which could not be afforded by the government and schools in the early 1980s. The recommendation of English as the dominant medium of instruction in the second phase can be considered as the extension of the emphasis on input in the second phase. CBEI came as a solution to the constraints such as lack of multiple copies of reading materials in the first phase. The emphasis on written input, i.e., reading, shifted to the emphasis on oral input, i.e., English as the medium of instruction. This effort, however, was impeded by an extensive range of contextual and social factor, such as a lack of supportive sociolinguistic environment for English medium instruction and a lack of teachers who were capable of bilingual education (Hu & Alsagoff, 2010). On second thoughts, even if CBEI can be implemented smoothly in mainland China, English medium instruction is not an optimal way of providing input to the students. As shown in Table 1.14 below, a comparison of the Schonell et al., (1956) count of spoken English and the Kucera and Francis (1967) count of written English indicates that written English contains twice as many word types as does spoken English (cited in Coady & Huckin, 1997), which suggest that oral input does not hold a candle to written input in terms of lexical density and variety.

Table 1. 14 Frequency of rare words in various sources

| | Rare Words per 1,000 |
|---|----------------------|
| I. Printed texts | |
| Abstracts of scientific articles | 128.0 |
| Newspapers | 68.3 |
| Popular magazines | 65.7 |
| Adult books | 52.7 |
| Children's books | 30.9 |
| Preschool books | 16.3 |
| II. Television texts | |
| Prime time adult shows | 22.7 |
| Prime time children's shows | 20.2 |
| Mr. Rogers and Sesame Street | 2.0 |
| III. Adult speech | |
| Expert witness testimony | 28.4 |
| College graduates talk to friends/spouses | 17.3 |

(Graves, 2006, p. 41)

In the third phase, the emphasis was shifted from input to output as evidenced by the boom of communicative language teaching, probably because of the unsatisfactory result of the effort to present written input in the first phase and the effort to integrate oral input in the second phase. The shift from input to output did not prove a panacea to all English language teaching problems in the Chinese context. Communicative language teaching failed to make the expected impact on English language teaching in China. Although a number of factors may account for the lower success of communicative language teaching, the most importance implication is that input is the basis of output. Without input, output did not produce the desired results. Thus in the fourth phase, the emphasis on either input or output is weakened. The syllabuses were replaced by a curriculum standard, which prescribes clear requirements on four language skills, i.e., listening, speaking, reading and writing, without specifying how to attain them.

In summary, various initiatives to reform English language teaching in mainland China failed to produce desired results because of a comprehensive range of social and educational factors. The core thread running through these factors is a lack of input. The lack of input has been a persistent problem throughout the introduction and description of the national syllabuses and curriculum standards although for different reasons.

As a variable in second and foreign language learning and acquisition and as an area of investigation, the concept of *input* has had an interesting and varied

history. In the behaviorist view of learning, input was a primary external mechanism, but not a casual variable itself (Gass & Selinker, 2008). In a cognitive view of second language acquisition, input is seen as the trigger that sets off the internal mechanism (Krashen, 1981). In language teaching, input means the linguistic material to which the learners are exposed, i.e., the learners' experience of the target language in all its manifestations (Smith, 1993). In acquisition, input has been identified as a crucial factor. Krashen states that language input plays an identical part with Language Acquisition Device (Jia, 1996).

The Input Theory proposed by Krashen advocates that input is primary and output is secondary in the early stage of second language acquisition (Sun, 2009). Output has numerous synonymous, such as "producing language" (Swain, 1995, p. 125), "language performance" (Swain, 1995, p. 125), "using the language", "speaking or writing" (Swain, 1995, p. 127), and "production or use" (Swain & Lapkin, 1995, p. 371). As introduced in the previous section, speaking and writing ability and communicative language competence can be operationalized as "output". The teaching practice and the various initiatives to reform English language teaching have shown that input is prior to output. Without input, there is no output. As summarized in Renandya (2007), the variables that may account for the unsatisfactory achievements in a second or foreign language include: inappropriate teaching methods, outdated textbooks, large class size, unqualified teachers, reluctant learners, lack of opportunity to use the target language and lack of input. Among the variables noted above, the lack of input is considered to be most closely related to the unsatisfactory achievements in a second or foreign language.

1.4 Finding a solution to the problem

The problem with input is that the researchers disagree on the strength of the importance of input although they agree on the importance of input in second language learning and acquisition. Is input the primary variable or is it one of a number of variables responsible for second language acquisition, such as motivation, opportunities for output, individual differences, and negative effect? In addition, little is known about the density and duration of input relative to the

acquisition of a second language. For many researchers, input and its role in second language acquisition remain complicated issues within the field. However, if we accept the premise that attention to comprehensible input is central to making language gains, then the role of input in the second and foreign language classroom must be a primary concern for language teachers.

In teaching practice, one of the biggest problems with input is that input in the classroom can be quite limited, which is particularly problematic in traditional classrooms where most of the activities are oriented to explicit instruction. An optimal vehicle for input is reading, which can be done both in and out of classroom. However, there are common reasons for the non-popularity of reading in these thirty years. One explanation for the non-popularity of reading, at least in the Chinese context, is that reading fluency is often not the ultimate goal of curriculum; instead, the goal is the development of language skills, as shown in the syllabuses in different phases.

The second explanation is that teachers are not willing to rethink how to teach students to read, because they lack the linguistic and reading skills to answer various questions that students may raise in reading. A third explanation is that teaching reading in class may subject teachers to criticism if students do not prepare for high-stakes exams in class. Both administrators and teachers have a vision of “teachers teaching something to students”. Extensive reading takes away that vision and teachers, in effect, feel disempowered.

Besides the assumed explanations for the non-popularity of reading, there are practical constraints on reading, among which lack of instructional time is the biggest. It is generally believed that learners need at least two hundred hours’ exposure to a language to obtain consistent and measurable achievement (Nunan, 2002). However, this two hundred hours’ exposure requirement cannot be met in the Chinese educational system (Yang, 2006).

Take the case of Shanghai as an example. Shanghai localizes NECS by supplementing NECS with specific requirements. As illustrated in Table 1.15, the time allocated for the English curriculum in twelve years is 1,518 sessions of

40-45 minutes, which can be converted into 1,012 to 1,138 hours, around 85 hours per year, 14 minutes per day.

Table 1. 15 Time allocated for the English curriculum in Shanghai primary and secondary schools

| Phase | Grade | Basic courses | | Development courses | Research courses |
|--------------------|-------|---------------|-------|---|------------------|
| | | Weekly | Total | | |
| Primary School | 1 | 3 | 102 | Schools are advised to arrange English extracurricular activities every week. | |
| | 2 | 3 | 102 | | |
| | 3 | 4 | 136 | | |
| | 4 | 5 | 170 | | |
| | 5 | 5 | 170 | | |
| Sub-total | / | / | 680 | | |
| Junior High School | 6 | 4 | 136 | Schools are advised to arrange development and research English extracurricular activities according to their own schedule. | |
| | 7 | 4 | 136 | | |
| | 8 | 4 | 136 | | |
| | 9 | 4 | 136 | | |
| Sub-total | / | / | 544 | | |
| Senior High School | 10 | 3 | 102 | No more than 170 sessions in 3 years. The schools arrange the research courses on students' request. | |
| | 11 | 3 | 102 | | |
| | 12 | 3 | 90 | | |
| Sub-total | / | / | 294 | | |
| Total | / | / | 1,518 | | / |

To supplement the insufficient instructional time in class and increase students' exposure to English, the Shanghai Curriculum Standards specifies the amount of extracurricular reading and audio and visual input (see Table 1.16 below). Based on the statistics, it is not difficult for us to work out the time and exposure to English. The required total amount of extracurricular audio and visual input is 280 hours in 12 years at most, which is equal to 4 minutes per day. The total reading amount for 12 years is 600,000 words at most, which is equal to 137 words per day. Senior high school students in Shanghai are required to read 250,000 running words in three years. A simple calculation will tell us that they need to read about 230 words per day if they keep reading every day. If they read every other day, 500 words each time is enough.

Table 1. 16 Amount of extracurricular input required by the Shanghai Curriculum Standards

| School stage | Level | Amount of extracurricular audio and visual input (No. of hours) | | Amount of extracurricular reading (No. of words) | |
|--------------------|----------|---|--------------|--|---------------------|
| | | Net | Accumulative | Net | Accumulative |
| Primary school | Basic | 60 | 60 | 800,000 | 800,000 |
| | Advanced | 20 | 80 | 200,000 | 1,000,000 |
| Junior high school | Basic | 90 | 150-170 | 1,800,000 | 2,600,000-2,800,000 |
| | Advanced | 10 | 160-180 | 200,000 | 2,800,000-3,000,000 |
| Senior high school | Basic | 90 | 240-270 | 2,500,000 | 5,100,000-5,500,000 |
| | Advanced | 10 | 250-280 | 500,000 | 5,600,000-6,000,000 |

Although the situation seems unpromising based on the calculation, the complementary of amount of extracurricular input into the curriculum standard shows that the syllabus designers are aware of the importance of input. At the same time, the NMET (Shanghai Version), the most influential factor affecting English language teaching in Shanghai, is sending the same signal that we should increase input in English language teaching.

The National Matriculation English Test (NMET) in China is the only test of English for university admission in the whole country (see Cheng, 2008; Li, 1990; Qi, 2005, for its prominent role in the Chinese educational system). Shanghai was authorized to develop its own version of English matriculation test, namely National Matriculation English Test (Shanghai Version) (NMETs) in 1985 (for the description of the national version of NMET, see Cheng & Qi, 1996). Although the name and part of the test formats were changed, its nature and influence remains the same.

Table 1.17, which summarizes the composition of language proficiency areas and language knowledge components in the NMET (Shanghai Version), illustrates the change of the proportion of reading in the NMET (Shanghai Version) since 2004. As illustrated in Table 1.17, the committee of the NMET (Shanghai Version) has shown a tendency to increase the proportion of reading in the exam. Since 2005, the proportion of marks that reading accounts for has been the biggest among all the subsections. In 2008, cloze was not tested as an independent subsection but was integrated into reading comprehension. Thus reading has accounted for one third of the total marks. In addition to this steady increase, the committee has hinted at further increasing the marks allocated to reading in this high-stakes exam.

Another progress was made in the test format. Until 2005, multiple-choice and blank-filling items constitute the vast majority (at least 85%) of the items in Paper One, and for the reading section, the percentage was 100%. In 2008, a new test format, Question and Answer (Q&A), was added to the reading section. As shown in Table 1.18 below, the latest reading section is divided into two parts: cloze and reading comprehension. The cloze section has 15 items in a

Table 1. 17 Composition of the language proficiency areas and language knowledge components in the 2004/2005/2008 NMET (Shanghai Version)

| Year | | 2004 | | 2005 | | 2008 | |
|----------|-------------|-------|-------|-------|-------|-------|-------|
| Paper | Subsections | Marks | % | Marks | % | Marks | % |
| Paper I | Listening | 30 | 20.0 | 30 | 20.0 | 30 | 20.0 |
| | Grammar | 20 | 13.3 | 16 | 10.7 | 16 | 10.7 |
| | Vocabulary | 10 | 6.7 | 4 | 2.7 | 9 | 6.0 |
| | Cloze | 20 | 13.3 | 20 | 13.3 | N/A | N/A |
| | Reading | 30 | 20.0 | 35 | 23.3 | 50 | 33.3 |
| Paper II | Translation | 15 | 10.0 | 20 | 13.3 | 20 | 13.3 |
| | Writing | 25 | 16.7 | 25 | 16.7 | 25 | 16.7 |
| Total | / | 150 | 100.0 | 150 | 100.0 | 150 | 100.0 |

four-option MC format. In the cloze section, fifteen words or phrases are taken away from a passage of 300 to 500 words. Candidates are required to fill in the blank with the most appropriate choices from the provided four options. In the reading comprehension section, there are five passages. The first three passages are for comprehension questions. After reading each of the first three passages, candidates are requested to answer three to four questions in a four-option multiple-choice format according to their understanding of their passage. The fourth passage is divided into five paragraphs. Candidates are required to choose five suitable sub-headings for the five paragraphs, one for each paragraph, from the six options provided. For the last passage, the candidates are requested to answer four questions based on the passage in their own words. No possible answers will be presented in a multiple-choice format (see Table 1.18 below).

Table 1. 18 Composition of the latest NMET(s)

| | Sections | Subsections | Item type | No. of Items | Marks | % | % |
|-----------|----------------------|----------------------|----------------|--------------|-------|-------|-------|
| Paper One | Listening | Short conversations | MC | 10 | 10 | 6.7 | 20.0 |
| | | Passages | MC | 6 | 12 | 8.0 | |
| | | Longer conversations | Blank filling | 8 | 8 | 5.3 | |
| | Grammar & Vocabulary | Grammar | MC | 16 | 16 | 10.7 | 16.7 |
| | | Vocabulary | MC | 9 | 9 | 6.0 | |
| | Reading | Cloze | MC | 15 | 15 | 10.0 | 33.3 |
| | | Reading | MC/ Q&A | 20 | 35 | 23.3 | |
| Paper Two | Writing | Translation | Transla tion | 5 | 20 | 13.3 | 30.0 |
| | | Guided Writing | Guided Writing | 1 | 25 | 16.7 | |
| Total | / | / | / | 89+1 | 150 | 100.0 | 100.0 |

The consistency between testing and curriculum standards may in fact provide the conditions for the development of extensive reading in mainland China. The

effect of extensive reading has been recognized world-wide in the past two decades, as evidenced by the existing body of research literature (e.g., Alessi & Dwyer, 2008; Al-Homoud & Schmitt, 2009; Ariyanto, 2009; Asraf & Ahmad, 2003; de Morgado, 2009; Elley, 1991; Furukawa, 2008; Gee, 1999; Greenberg, et al. 2006; Hayashi, 1999; Iwahori, 2008; Krashen, 2004a; Renandya & Jacobs, 2002; Rodrigo, et al., 2007; Roszell, 2007; Sin, 2007). Research shows that extensive reading is not in opposition to the Grammar-Translation Method that has long dominated mainland English language classrooms but is a feasible and quality complement to it (Renandya, 2007), as well as a useful complement to a communicative-approach based language class (Powell, 2005). Since the benefits of extensive reading have been extensively discussed in the research literature (see Section 2.2 for detailed introduction of benefits of extensive reading) and it is not incompatible with a variety of teaching methods, it is reasonable to suggest that extensive reading be incorporated into high school English language lessons to motivate students to read and enhance their English language proficiency.

To summarize, extensive reading is an optimal alternative that meets the following demands:

1. It is not in conflict with the traditional Chinese culture and syllabus, nor curriculum and teaching material.
2. It increases students' exposure to English both in and outside the classroom.
3. It promotes students' improvement both in the proficiency areas that the NMET lays emphasis on as well as their all-round development in English.
4. It can be achieved both in well- and poor-resourced schools.
5. It has been supported by research literature.

1.5 Purpose of this study

It is in recognition of the restrictions imposed on the initiatives to reform English language teaching and the results obtained that syllabuses are changed into curriculum standards, which leaves a wider space for teaching practice. Of particular note is the increasing emphasis on input, in both oral and written forms, which is the strongest justification for extensive reading in the Chinese context. However, to date there is a dearth of extensive reading studies in this context,

and there is no recognized scheme of extensive reading in mainland China.

Thus this research undertaking aims to provide an extensive reading program through which students are immersed in an input-rich environment. Despite successful research and a growing interest in extensive reading in many Asian countries, such as Hong Kong, Japan, and Singapore (see Section 2.5 for the introduction of extensive reading schemes in Hong Kong, Japan and Singapore), extensive reading has not received as much attention as it might in the English language teaching circle in Shanghai. Attempting to establish the benefits of extensive reading in Shanghai senior high schools is a valuable and researchable enterprise.

1.6 Significance of this study

In recent years, there have been some studies in China on improving students' English language proficiency. However, most of the studies applied to English language teaching in college or higher education institutions rather than secondary schools, and few studies touched the core of English language teaching in China—its failure to provide extensive exposure to the target language. Not surprisingly, the benefits of large exposure to the target language have not been achieved and students' English language proficiency has not significantly improved.

In the past decades, most of the studies on the significance of extensive reading show that it can effectively improve ESL learners' English language proficiency (Elley & Mangubhai, 1983). However, this body of research does not apply to Chinese secondary schools, because English language teaching and learning in this context is quite different from English language teaching and learning in an ESL context in terms of learners' exposure to English through listening, reading and writing inside and outside class. The Chinese secondary students' exposure to English may not as extensive as that of many ESL learners elsewhere. Compared with the Chinese EFL learners, ESL learners may live in English-speaking environments and their exposure to the target language is also much larger than that of Chinese secondary students. Therefore we cannot say that what is true in EFL and ESL teaching and learning abroad is certain to be

true in Chinese secondary schools.

This study focuses on the secondary school students in Shanghai, mainland China. Its purpose is to see the effect of extensive reading on a specific group of EFL learners. If this study is successful, the outcomes of the present study will be used to recommend feasible changes or modifications in the current practice in English language teaching in Chinese secondary schools.

The following three points summarized in Hu (2002) explains why extensive reading in secondary schools deserves particular research attention. First, English language teaching and learning at this level has experienced many shifts and changes since 1978, as evidenced by the frequently updated national syllabuses and curriculum standards which were detailed in Section 1.2. Second, English language teaching at secondary school level is of paramount importance to Chinese EFL learners because the vast majority of them begin to learn English in secondary school. English language teaching in secondary school helps them build knowledge foundation and cultivate good habits for further study of English. Third, the number of learners influenced by English language teaching at secondary level is larger than that at primary and university level. These three reasons make English language teaching at the forefront of the educational reforms.

1.7 Research questions

The present study attempts to bridge this gap and examines whether extensive reading programs can be successfully implemented with EFL secondary school students in Shanghai by looking at the following questions:

1. Does extensive reading result in improvement in Shanghai secondary school students' English language proficiency? Specifically, does extensive reading positively influence the students in terms of language knowledge, namely, grammar and vocabulary, receptive skills, namely reading and listening, or productive skills, namely writing (and translation in the present study)? If so, in what ways?
2. How can extensive reading be more successfully implemented in the Chinese context? Which type of extensive reading program is more effective,

one that is integrated as part of the class curriculum, or one that is an extracurricular activity?

3. Which form of extensive reading implementation may be more effective in enhancing a particular skill?

1.8 Hypotheses

Our hypotheses are as follows:

1. Extensive reading will result in improvement in Shanghai secondary school students' overall English language proficiency.
2. Extensive reading will positively influence the students in terms of language knowledge, receptive skills and productive skills.
3. The effect of extensive reading will be greater in the receptive skills than the productive skills.

2 Literature Review

2.1 Characteristics of extensive reading

2.1.1 Definitions of extensive reading

The history of extensive reading can be dated back to medieval times, but Harold Palmer was credited by Kelly (1969) as the first to coin the term “extensive” in 1917. Previous studies on extensive reading have focused on its definition, how it affects language teaching and learning, and how it can be implemented optimally in the classroom (Hill, 2008).

Among the many issues under investigation, the definition is central as it is closely related to many other issues such as implementation and evaluation. There are numerous definitions of extensive reading by many researchers, and in essence these definitions can be grouped into three categories: extensive reading as a reading style, extensive reading as an input medium and extensive reading as a teaching approach.

Extensive reading as a reading style

Arnold (2009) calls extensive reading a reading style, comparable to skimming and scanning. “Today, in language-teaching terms, extensive reading is recognized as one of the four styles or ways of reading, the other three being skimming, scanning, and intensive reading” (Day & Bamford, 1998, p.6). Researchers defining extensive reading from this perspective focus more on the reading process itself rather than what language learning outcomes it may yield.

A distinction between extensive reading and intensive reading suggests that extensive reading is reading a whole story or chapter for general comprehension after class, whereas intensive reading means reading two to four pages for detailed linguistic analysis in order to teach grammar and vocabulary (Brooks, 1960). Yet, King (1969) suggests that this distinction is difficult to maintain since both extensive reading and intensive reading are done with the same text, and in practice few readers read with a timer in hand to distinguish his or her reading from skimming, scanning and intensive reading. There is indeed a distinction between intensive and extensive reading, but the boundary between the two is

usually burred.

The value of defining extensive reading as a style instead of an approach is that extensive reading can be associated with incidental vocabulary acquisition, either acquiring new lexical items or expanding knowledge of lexical items that have been known to the readers, without focusing on form (e.g., Schmitt, 1998, 2000; Stahl & Nagy, 2006). An area of inquiry directly related to defining extensive reading as a reading style is how many exposures are needed for acquisition to occur (e.g., Waring & Takaki, 2003). Another area of inquiry closely related to the previous area of inquiry is the condition under which satisfactory text comprehension will occur, i.e., what percentage mastery of the running words in the text will ensure smooth and successful comprehension (e.g., Hu & Nation, 2000; Nation, 2001; Pulido, 2009). This area of inquiry can be further extended to branching topics, such as how many words there are in English (e.g., Carroll, et al., 1971; Nation, 2001), how many words L1 students know (e.g., Nagy & Anderson, 1984; Stahl, 2005), how many words L2 students know (e.g., Nation & Waring, 1997), and accordingly the vocabulary size needed to handle texts intended for L1 learners, L2 learners, academic, technical and general purposes, etc (e.g., Hirsh & Nation, 1992; Laufer, 1992; Nation, 2006), and finally appeal for using simplified texts or graded readers (e.g., Blau, 1982; Claridge, 2005; Davis, 1984; Honeyfield, 1977; Leow, 1993, 1997, 2009; Nation & Wang, 1999; Widdowson, 1976; see also Section 2.3 for discussion on graded readers).

Extensive reading as an input medium

The stress on the amount of reading is a general concern among extensive reading researchers and reflected in many of its proposed definitions, for example, “reading in quantity and in order to gain general understanding of what is read” (Richards & Schmitt, 2002, p. 193), and “exposing learners to large quantities of material within their linguistic competence” (Grabe & Stoller, 2002, p. 259). For the purpose of our discussion, this concern has brought another perspective in defining extensive reading: extensive reading as an input medium.

Within the frame of extensive reading as an input medium, the major disagreement among the researchers lies in the level of difficulty of the reading

material. As proposed in Grabe and Stoller's and Day and Bamford's definitions, the level of difficulty is preferably "within their linguistic competence", based on which extensive reading is "reading gain without reading pain" (Day & Bamford, 1998, p. 121). This is known as the *i-1* claim, i.e., the difficulty level of reading material should be slightly lower than the current level of the readers (*i*). The biggest challenge to the *i-1* claim comes from Krashen (2004a). He argues that "Enjoyment does not guarantee language acquisition" (p. 28), advocating that input needs to be at the level of the acquirer (*i*) with some items just beyond his/her grasp (+1). Krashen encourages the readers to move a step further beyond their level, because it is indeed exactly one goal of extensive reading: one sign of the improvement of a reader's reading ability is that readers can read more difficult reading materials. The problem with Krashen's *i+1* claim is that in practice it is difficult to grasp the subtle balance between pleasure and difficulty. It is contentious whether pure pleasure will bring no progress or too much difficulty will ruin pleasure, leading the readers to lose their interest or motivation for extensive reading. It is not easy to justify the old saying "No pain, no gain" in an extensive reading program. A practical solution to this controversy is that if the reading material is easy, readers are certain to read faster. The vicious circle of the weak readers presented in Nuttall (1996) is that the weak readers do not understand what they read because of lack of vocabulary, so they read slowly. They do not enjoy reading. Consequently they do not read much. As a result the vicious circle repeats. One of the direct consequences of slow reading is that the readers will be exposed to much less amount of comprehensible input, which is central to language acquisition (see Section 4.6.1 for the detailed discussion on the value of fast reading).

Extensive reading as a teaching approach

Day and Bamford (1998) identify extensive reading as a teaching approach, proposing that "Extensive reading is an approach to the teaching and learning of second language reading in which learners read large quantities of books and other materials that are well within their linguistic competence" (p. xiii).

Although no time-scale is actually mentioned by Day and Bamford in their taxonomy of what they consider the primary characteristics of extensive reading,

it must be assumed that extensive reading, like “reading extensively”, applies to longer rather than shorter time periods. This requirement may, in turn, raise practical difficulties, as the longer an extensive reading programme runs, the greater the chance of extraneous variables impinging on the program. This has been found to be particularly the case in Africa where it was difficult to maintain professional collaboration with local teachers who were poorly paid and had to hold multiple jobs (e.g., Davidson & Williams, 2005; Williams, 2007).

It is very difficult to control variables in the normal classroom for the length of time required to be able to lay any credible claim to having investigated experimentally the effects of extensive reading, which partially explains why so few longitudinal extensive reading programs have been reported. The second potential threat to a long-term extensive reading program is that it is difficult to gain high or professional levels of cooperation from others, i.e., to have the local teachers teach one group for a long period of time; even if this high level of cooperation can be achieved, it still puts the program at risk because the local teachers may not have been professionally trained for extensive reading. In actual extensive reading programs, the first solution is that the researchers teach both experimental and comparison groups themselves (e.g., Bell, 2001). An alternative solution is that the researchers use a previously evaluated comparable group as a pseudo control group (e.g., Lai, 1993b). Both of these two solutions may carry potential threats to the research design. This is often the case in small-scale extensive reading studies. Large-scale extensive reading programs, especially government-funded extensive reading programs, have to require high levels of cooperation from classroom teachers. In this situation, training of the local classroom teachers is of particular importance, and proper training proves to be very effective in ensuring successful implementation, as reported in the Reading and English Acquisition Program (REAP) in Singapore (see Section 2.5.2 for details of REAP).

Because of the practical difficulties involved, consequently, although very dynamic in research literature in recent decades, extensive reading is rarely claimed as an “approach” for language teaching, i.e., the Extensive Reading Approach. Extensive reading so far has not obtained a place in the language

teaching approach and method family (see Richards & Rodgers, 2001, for main-trend and well-recognized approaches and methods in language teaching in the 20th century).

Summary

How we define extensive reading, especially extensive reading as an input medium and extensive reading as a teaching approach will greatly influence the way we implement extensive reading in the classroom. A major challenge involved in defining extensive reading as an input medium is whether to attribute the benefits of extensive reading to the exposure to the target language or the explicit instruction in the classroom. This is particularly problematic in an ESL context, as the participants in both experimental and control groups should not have additional periods of L2 exposure. This challenge can be largely avoided in an EFL context, where EFL students have little or no exposure to English outside the classroom.

The challenge involved in using extensive reading as a teaching approach in an EFL context is the lack of face-validity of “reading only” as a teaching approach. For example, as introduced in Chapter 1, the administration, the teachers and the students are not happy with the teachers “teaching nothing” in class. Consequently certain compromises and modifications have to be made to implement extensive reading smoothly in the Chinese context. We will discuss possible compromises and modifications in Chapter 3 and we will suggest more comprehensive recommendations for large-scale extensive reading programs in Chapter 6.

2.1.2 Agreed characteristics of extensive reading

Although there are different definitions of extensive reading, there are agreed characteristics of extensive reading. The characteristics inherent in extensive reading are: reading quantity and the various benefits it is supposed to bring to the readers.

Reading quantity

A core justification of extensive reading is that extensive reading generates a

virtuous circle, which means the readers will be better at reading by reading more, and they have more fun from reading by being better readers (Jensen, 1986). One of the core characteristics of extensive reading, as its name shows, is “extensive”, which suggests and request reading in quantity. So far there has been little agreement on what constitutes “reading in quantity”: what amount of reading can meet the basic requirement. Derived from different definitions of extensive reading, there are at least four perspectives of defining amount: absolute amount, reading speed, reading time and the ratio between extensive and intensive reading.

The optimal reading amount and even the counting units for recommend optimal reading amount vary across from different researchers. For example, by employing “page” as the counting unit, Matsumura (1987, cited in Susser & Robb, 1990) suggests one page one day on school days and three pages one day during holiday. By adopting “book” as the counting unit, Hill (1983) recommends a minimum of 36 graded readers every year, and Bright and McGregor (1970) encourage 60 books per year. Among these variations, the 60 books per year proposed by Bright and McGregor (1970) seems to win broader support, as Day and Bamford (2002) set a goal of one simplified reader per week for extensive reading programs. Nation and Wang (1999) also confirm the one-per-week standard to be the minimum needed to ensure that learners get the repeated exposure to the new words that make them memorable. However, one book per week is not the upper limit. It is not a perfect amount that enables a reader to make dramatic achievement through reading a book per week. As Flesch and Lass (1996) point out, the guiding principle for the amount of reading is “The more you read the better. A book a week is good, a book every other day is better, a book a day is still better.”

It must be pointed out that the counting units “*page*” and “*book*” are vague. No strict guideline has been issued to specify what constitutes “a page” or “a reader”. For example, whether a page with a picture counts the same as one without? Taking the Oxford Bookworms (OBW) series as an example, the word count of an OBW Level 1 book varies from 5,000 to 7,000, while that of an OBW Level 6 book is between 25,000 and 35,000. Following the train of logic, even if the

readers adhere to the one book per week guideline, the actual reading amount may vary enormously among readers, and in the long run the difference could accumulate to a huge figure.

Taking into account this possible scenario, Helgesen proposes a uniformed “page” system as follows by sorting out the various counting units and levels of difficulty:

Level 1: 1 page = .5 page

Level 2: 1 page = .75 page

Level 3: 1 page = 1 page

Levels 4 - 6: 1 page = 1.25 pages

(Helgesen, 1997, p. 32)

Helgesen has made an effort to quantify reading amount by associating readers’ level and the amount that they are expected to manage, which reveals that reading amount is not absolute but should be within the readers’ competence.

It has been suggested that a more objective counting unit would be the reading duration. Hafiz and Tudor (1989) suggest 60 hours over 3 months, and Newmark (1971) recommends two hours per week. The problem with this counting unit is that individual reading speed varies enormously. Some are able to read at 200 words per minute, while some cannot manage 60 words per minute. The result is that even if the reading time is guaranteed the actual reading amount in the long run may well vary hugely.

Thus “quantity of reading is not an absolute number of hours or pages but depends on teacher and student’ perceptions of how extensive reading differs from other reading classes” (Susser & Robb, 1990, p.166). Based on this notion, two more suggestions for reading quantity are made. Tanaka and Stapleton (2007) suggest a proportional relationship between extensive reading and the text book in terms of the number of words read, stating that extensive reading means reading three times as much as the number of words in the textbook. The other suggestion made by Williams (1986) proposes a proportional relationship between extensive reading and intensive reading in terms of time on reading: one

hour of extensive reading for one hour of intensive reading.

As can be concluded so far, no unified or mandatory standards or requirements for optimal reading amount have been generated from the research literature, but a feasible alternative put forward by Light (1970) is that the reading materials must be of sufficient length, which excludes the possibility that the teachers use the materials for in-class reading. Helgesen (2005) applies this rule by setting a 500-page per semester goal for his readers, stating that students would never manage it if they read the text word by word, line by line. As a whole, reading quantity is set beyond the capacity of intensive reading instruction, with extensive reading primarily an out-of-class activity (Waring, 1997).

The latest trend in quantifying reading amount is the word count of each reader. OBW presents the word account of each book on the back cover, which is a great benefit for teachers who would like to keep a more exact record of students' reading amount. Here it also should be pointed out that it is one thing to present the word count, but it is another thing to quantify the effort students have made to process the running words. It requires certainly different intelligence and knowledge to decode the most frequent 1000 words and the 1000 most specialized technical terms in texts.

Presently what can be concluded is that "*the more reading the better*" principle should be based on readers' own proficiency level and their own perception of what extensive reading is, although reading in quantity is a common feature of extensive reading.

Benefits of extensive reading

The second characteristic of extensive reading is the various benefits it is supposed to bring to the readers, which none of the extensive reading researchers would disagree with each other, no matter what perspective they take on extensive reading. In fact, this is where the strength of extensive reading lies.

The research literature has well documented a comprehensive range of benefits accruing from extensive reading (e.g., Day, 2004; Dupuy & Krashen, 1993;

Davis, 1995; Eskey, 2002; Grabe & Stoller, 1996; Hunt & Beglar, 2005; Iwahori, 2008; Laufer, 2003; Lee, 2006; Lee & Hsu, 2009; Pigada & Schmitt, 2006; Renandya, 2007; Tanaka & Stapleton, 2007; Webb, 2008; Yamashita, 2007), from improvement in reading itself to improvement in language knowledge and other skills, i.e., grammar and vocabulary, receptive skills, i.e., listening as well as productive language skills, i.e., writing and speaking. Specifically speaking, in addition to developing reading proficiency, extensive reading is “intended to develop good reading habits, to build up knowledge of vocabulary and structure (c.f. definition of extensive reading as a reading style), and to encourage a liking for reading” (Richards et al., 1992, p.133). Besides its effect on vocabulary, structure and interest in reading, extensive reading leads to “vocabulary recycling, sight word reading gains, reading speed gains, reading practice, and a habit of being a successful reader in the L2”, and accordingly it is a central rather than peripheral component in L2 reading instruction (c.f. definition of extensive reading as a component to teaching reading quality) (Grabe & Stoller, 2011, p. 122). Besides its role in language proficiency development and reading instruction, “extensive reading plays a key role in moving learners beyond the basic levels of proficiency required for basic communication so that they are well positioned to achieve personal and professional goals” (Horst, 2009, p.63).

In addition to the benefits in the target language, there is noted transfer of gains from the target language to other areas. Elley (1991) outlines a series of empirical studies focusing on the effects of extensive reading by means of “book-floods” on elementary school students’ second language acquisition in Fiji. Children in these studies learnt the target language incidentally through extensive reading, and in some studies, the benefits have been found to spread to other school subjects and vernacular. Students in Singapore who had taken the Extensive Reading and Acquisition Program (REAP) increased their chance of passing the public examination by 10% (Guo, 2007).

The problem is, although the benefits are considerable, the ways to categorize them differ. Davis (1995) classified the benefits into the following four interrelated headings: reading skills, language skills, personal growth and examinations. In the first book to focus specifically on extensive reading,

Extensive Reading in the Second Language Classroom (Day & Bamford, 1998), the benefits of extensive reading are presented as: results of extensive reading programs, second language ability, affect, vocabulary, linguistic competence and spelling, probably from the perspective of its implementation, and from a macro to micro level in terms of its benefits. Jacobs et al. (2007-the present) published an online annotated bibliography of works on extensive reading on the Internet, aiming at presenting a comprehensive picture of research on extensive reading. The framework they adopted classifies that benefits of extensive reading under the following headings: affect, learner autonomy, L2 learning, reading ability, grammar, vocabulary and writing.

The categorizations above are reasonable in their own right. As discussed above, Day and Bamford define extensive reading as an approach for teaching quality wherein it is logical to assume that extensive reading will result in improvement in second language ability and language competence. Davis coordinated a project to aid the improvement of English language skills in selected schools, thus it was not surprising that reading ability was a primary concern. It should be noted that, although the above and many other categorizations of benefits of extensive reading are reasonable in their own right, a universal categorization framework is still not available because of lack of agreement on the definition and the varied objectives of extensive reading programs. Some terms employed in the categorizations mentioned above are ambiguous. For example, there is no strict boundary line between competence and ability; similarly, spelling can also be part of vocabulary learning also since Nation (2001) has distinguished nine levels of knowing a word, with knowing the spelling of a word as the second level. The essence of these differences is that the experimental enquiry into the effectiveness of extensive reading has so far produced little more than a collection of somewhat disparate findings. Nor has any attempt been made to categorically link any such research findings with second language acquisition theory (Irvine, 2006). Consequently, extensive reading lacks coherent and research-based theory as has been introduced in Section 2.1.1.

Another difficulty in presenting the benefits of extensive reading is that despite the appealing benefits that extensive reading may bring with it, research has

shown that the benefits of extensive reading cannot be easily observed over a relatively short period of time (Grabe & Stoller, 2002). Although the benefits are considerable, it is difficult to quantify without longitudinal studies as the benefits may not be detected by immediate follow-up tests or assessments (Davis, 1995). Furthermore, Yamashita (2008) points out that extensive reading may generate a global effect on overall language proficiency in the long run; however, its effect on a particular language skill may not be equal in a relatively short period of time.

Thus the first challenge that extensive reading faces is that tiny incremental growth can hardly persuade curriculum designers and language teachers to implement extensive reading programs into their schools or incorporate in the curriculum; at the same time, extensive reading will disappoint those who hope to make dramatic achievements in a particular proficiency area.

The second challenge faced by extensive reading practitioners is that there is not yet a universal instrument tool to standardize the effect of extensive reading across different countries and regions. The reported benefits so far are established by self-developed exams or small-scale instrument tools. For the validity of the effect, and also the feasibility of extensive reading assessment, a highly reliable and practical assessment tool of extensive reading that is applicable and standardized across different countries and regions is urgently needed.

2.1.3 A less-discussed issue: reading quantity vs. reading quality

Another issue hidden from public attention by concern on reading quantity is reading quality. Some researchers have made their position on this clear i.e. that the watch words of an extensive reading program should be quantity and variety (e.g., Day & Hirsch, 2004).

The argument for reading quality comes from Krashen's Comprehension Hypothesis which postulates language acquisition happens when people understand what people tell and what they read (Krashen, 2004b). D. Hill (personal communication, April, 2010) proposes that reading quality is no less

important than quantity. Due to the lack of research literature on this topic, the researcher is hoping that the findings of the present study will help to provide evidence on the relationship between reading quantity and quality.

2.2 The research on extensive reading

As noted in the previous section, because of the lack of coherent and research-based theory of extensive reading, there is debate or disagreement on several core issues of extensive reading: definition, reading quantity, types of reading materials and organization of extensive reading programs. The loose and varied definitions and variations on the extensive reading theme have led to the diversity in the implementation of extensive reading programs. Earliest came Sustained Silent Reading, Interrupted Sustained Silent Reading, Free Voluntary Reading, etc. Since the late 1970s, many teachers and reading specialists have been recommending students engaging in Sustained Silent Reading, or Interrupted Sustained Silent Reading, most of which were done on the campus.

Sustained Silent Reading (SSR) is one of the variations on the extensive reading theme. It is carried out in school where students read silently in a given period of time every day. The assumption underlying SSR is that students learn to read by reading constantly (Cho & Krashen, 2001). SSR offers guidance for individual reading ability and interests, as well as providing opportunities for readers to explore, appreciate, and think through their readings, while simultaneously affording choice and pleasurable experiences in reading (Hsui, 1994). Empirical studies on SSR have been reporting students' gains in vocabulary knowledge (Nagy et al., 1985), reading level (Ozburn, 1995), reading rate (Dwyer & West, 1994), reading achievement (Coley, 1983) and more positive attitude towards reading (Wiesendanger & Bader, 1989). A general tendency observed from the SSR programs by a meta-analysis which collected ten SSR studies is that five SSP programs lasting longer than six months all reported positive outcomes whereas the remaining five SSR programs spanning less than six months unexceptionally reported no benefits (Wiesendanger & Birlem, 1984).

Another variation of SSR, the Uninterrupted Sustained Silent Reading (USSR), was employed by numerous schools in Singapore as a common strategy to

encourage reading. The practice of a school-based USSR was to gather all the students in some large space such as the school hall or canteen to read 20 to 30 minutes each session (Seow, 1999). Varying degrees of success were achieved by different schools. In some schools students seemed to enjoy their USSR sessions, while the USSR activities were not carried out successfully in some other schools for various reasons (Klapper, 1992).

To examine the varying degrees of success in implementation of USSR in different schools, Seow (1999) compared an USSR program with another reading program, the Extensive Reading Program (ERP). The significant similarities and differences revealed by the comparison suggested the following reasons for the lower success of USSR in some schools. First of all, although closely-related to each other, USSR and ERP are in fact different. An actual extensive reading session includes a sustained silent reading period but it may also involve some follow-up activities if done formally in the classroom, while in a considerable number of schools USSR was implemented in the form of an extensive reading program but without all the core components of a proper extensive reading program. This calls for a need for a clear understanding of USSR and ERP, and what should to be done in USSR and ERP respectively. Second, a great proportion of the time that should be devoted to reading was wasted in getting the students to be seated, and it was almost time for the next school session when the students finally settle down to read. This shows the strategy of centralization, i.e., gathering the students in a large space for reading, fails to achieve to its desired result, namely a global monitoring of the reading. The random flipping of the pages, the disinterested look on the readers' face and even the competition among the students for more physical space were frequently observed. These observations suggest that USSR is more effective with students in small groups and in an environment which is more conducive to reading. The third reason that may explain the lower success of USSR is the lengthy duration. Schools should have realized that students need to be promoted to independent self-reading when a habit of sustained silent reading has been instilled in them by USSR. They should be encouraged to visit the school library and read whenever they are available. Thus based on the above-noted reasons, it is suggested that USSR serves as a prelude to an extensive program, which is to say, the ultimate purpose

of USSR is to prepare the students for more independent extensive reading (Seow, 1999).

Free Voluntary Reading (FVR) is another variation of the extensive reading theme coined by Krashen (1989), which aims at readers' overall development in the target language. Krashen (2004a) believes the most effective tool to help eliminate illiteracy is FVR. FVR is very simple. Students can read anything they want. If they don't like a book, they opt for another one. Students also have no book reports to write, no questions to answer, and no looking up unknown vocabulary (Krashen, 2004a). By using FVR, Shin (2001) conducted a summer school study where students of limited English language proficiency were given approximately two hours a day for self-selected reading. Another 45 minutes each day was designated for assigned reading. In the post-program Altos test, the students gained five months of reading comprehension in a six-week period according to the test results. On another test, the Nelson-Denny Reading Test, the students gained over a year. As shown here, research shows that students can directly benefit from FVR.

More empirical studies show that FVR can not only be a way to improve reading comprehension, it also helps with vocabulary. MacDonnell (2004) detailed a project to help students read for pleasure and gain vocabulary recognition in which students in sixth and seventh grades begin the year with a six-week reading for pleasure program. Students choose a book they wanted to read and then finished it if they liked it. If they did not like it, they chose another book. Book reports were not required, but if a student liked a book, they were required to give a short recommendation to their peers. This resulted in more circulation of the books that were promoted by the other students. Peer recommendation has been shown by researchers as a great way to promote reading among students (Wendelin & Zinc, 1983).

Since the 1980s, the effect of flooding classrooms with books has been gaining popularity, and the Book Flood approach became a widespread concern about the teaching of English. In this decade, one of the most notable extensive reading programs, and also the first in a series of large-scale research studies on the

benefits of extensive reading and book-flood curricula in a number of countries around the world, was the Book Flood Project in Fiji. This project conducted by Elley and Mangubhai (1981) was set up in eight rural schools in Fiji in 1980. The hypothesis of this project was that the students would learn English more efficiently through being exposed to the influx of comprehensible input in that language (Elley, 1987).

This project was a pioneer in establishing two experimental groups. The students participating in this project were divided into two groups: half of the students were to participate in the Shared Reading Method, and the second half were to take part in the Sustained Silent Reading Approach. These two groups formed the Book Flood Group as the experimental group, and the rest of the students who did not participate in this project, as a whole, formed the Tate Program Group as the control group for comparison. The time spent on language was standardized. All of the approaches took up 30 minutes every day. No extra time was spent on any group. Reading material supplied to the Shared Reading Group and Silent Reading Group were 250 books each. The Control Group adhered to the normal syllabus and curriculum, and no extra books were provided.

The shared reading method was conducted in several consecutive sessions. In the first reading session, the teacher chooses a high-interest story with appropriate language and illustrations and introduced the pupils to a “sharing experience”, similar to that of a bed-time story. Discussion was encouraged about pictures, the likely contents, and a few new words. The teacher then read all, or some of it, to the class. To ensure that all of the group can see the text and illustrations, the text was frequently “blown-up” or rewritten in the form of a giant book, with suitably-sized illustrations. During the second or third readings, children were encouraged to join and read easier sections with the teacher, who continued to encourage discussion about the context of the books. Emphasis was placed on prediction and confirmation of events in the story, so that children were constantly striving for meaning. If children enjoyed the experience, they would want to read it often, in the class group, in small groups, in pairs, or as individuals. The intention was that they master the language of the book, with a minimum of pressure and strain. Follow-up activities included role play,

vocabulary learning, art work and writing activities. The origin of these activities is always determined by the story, not by any pre-determined system of the proper sequence of structures and vocabulary to follow. Theoretically, the essence of the shared book method is that learning occurs at the point of interest, rather than in accordance with a carefully graded linguistic pattern. In the Silent Reading Group, the main role of the teachers was to display the books attractively, to read them aloud regularly, and to spend 20 to 30 minutes each day in sustained silent reading, with books of the children's own choice. A definite period was set aside every day for reading, during which time the teachers set a good example by reading also. No book reports were required, and no written exercises were performed. The children read for enjoyment and practice. The teachers and students in the control group followed their normal curriculum in English. Students had two 15-minute oral English lessons each day in which new structures were systematically introduced, in appropriate classroom situations, with repeated drills, variations, substitution tables, etc. Reading was taught through graded readers or the Fiji Ministry's text. Both series of books were graded and activities were intended primarily to provide practice in order to consolidate the structure and vocabulary taught in the oral lessons.

A comprehensive range of language tests was administered in the following two years. The two reading groups significantly outperformed the control group by showing much greater improvement in reading, writing, listening comprehension, speaking and vocabulary. The general trend observed was that the benefits of extensive reading first emerged from the receptive skills, i.e., reading and listening, and later from writing, a productive skill, and vocabulary, a language knowledge component, and finally and more surprisingly, from other school subjects such as science and math. In general, the effect of extensive reading on receptive skills was greater than that on productive skills. Findings of the Book Flood Project in Fiji was consistent with the notion of Yamashita (2008) noted in Section 2.1.2 that extensive reading may generate a positive effect on overall second language proficiency in the long run, but its effect on a particular language skill may not be equal in a relatively short period of time.

Since 1990, additional evidence has accumulated to support the benefits of

extensive reading. In 1995, a “The Book in Schools” project was implemented in Sri Lanka with the assistance of the International Book Development. The local languages in Sri Lanka are Sinhalese and Tamil, and English is taught as a foreign language since Grade 3 in primary school. “The Book in Schools” project was intended for the Grade 4 and 5 pupils, one or two years after they began learning English. In all, thirty schools were selected for this project, of which twenty were in the experimental group and the remaining ten served as the control group. The control group spent all the forty minutes on the normal curriculum and textbooks, while the experimental group allocated fifteen to twenty minutes on shared reading of the books and spent the remaining time on the normal curriculum and textbooks (Elley & Foster, 1996). The post-project was carried out in January 1996, six months after the project began. The experimental group demonstrated substantial gains in vocabulary and reading comprehension by obtaining scores three times as high as those of the control group. Gains were also found on listening and writing, although less spectacular than gains on vocabulary and reading (Elley, 2000).

It must be pointed there that the Book Flood Project in Fiji and the Book in Schools Project in Sri Lanka are not the only two noticeable extensive reading programs before the new millennium. The Hong Kong Extensive Reading Scheme (HKERS) (1991-now), the largest-scale government-funded extensive reading project that has ever been undertaken, and the Reading and English Acquisition Program (REAP) (1985-1989) in Singapore, one of the most successful extensive reading programs around the world, are also important milestones in the extensive reading history. These two programs are detailed in Sections 2.5.1 and 2.5.2 respectively.

Entering the new millennium, growing research has accumulated to support the benefits of extensive reading. In addition to a growing number of empirical studies showing the effectiveness of extensive reading (e.g., Guthrie et al., 2004; Horst, 2005; Perfetti, 2007; Pigada & Schmitt, 2006; Tanaka & Stapleton, 2007), some of the most recent research on extensive reading has been updating the findings from the previous research on some topics. For instance, in the 1980s research on incidental vocabulary learning from extensive reading showed that 5%

to 10% new words were learned from extensive reading (Nagy, et al., 1985). Over time, research has updated this percentage to 5% to 15% (Stahl & Nagy, 2006), and the most recent finding is the readers pick up 35% of the new words in extensive reading (Horst, 2009).

It is also important to note that extensive reading began to attract research attention in mainland China in the new century. Research undertakings on extensive reading have been conducted at both secondary school level (e.g., Li, 2007; Li, 2007; Li, 2007; Li, 2008; Liu, 2011; Peng, 2004; Wang, 2008; Yao, 2005; Zhang, 2012; Zhang, 2003; Zhao, 2006; Zhu, 2011) and tertiary level (e.g., Gao, 2008; Qin, 2008; Ren, 2010; Zhang, 2005). These studies carried out all over China not only show the teachers' growing awareness of the importance of input, but also inspiringly demonstrate the feasibility and effectiveness of extensive reading in this context.

To summarize, research on extensive reading since the 1970s not only has generated a gamut of benefits of extensive reading, but also witnesses its growing popularity and actual implementation all over the world, including in third-world countries like China and Sri Lanka. The research evidence from the existing body of literature serves as hard backup and strong support of extensive reading.

Another note to make here is that the review of the research on extensive reading in this section aimed to give a brief overview of its history and most recent development. It does not mean that the research literature concerning extensive reading is limited to what has been noted in this section, or only large-scale extensive reading programs are worth mentioning. The research literature relevant to the research concerns in the present study (see Sections 1.7 and 1.8 for the research questions and hypotheses in the present study) will be reviewed and presented in corresponding sections in Chapter 4 with an aim to interpret the test results and address the research questions, i.e., extensive reading and vocabulary acquisition in Section 4.4.1, extensive reading and grammar acquisition in Section 4.4.2, extensive reading and reading proficiency in Sections 4.5.1 and 4.5.2, extensive reading and listening proficiency in Section

4.5.3, and extensive reading and writing proficiency in Sections 4.6.1 and 4.6.2.

2.3 Reading materials

Reading undeniably and incontrovertibly involves two necessary elements: a reader and a text (see Section 1.1 for detailed introduction of the subcomponents of these two elements). Our discussion thus is organized around readers and texts, and the interaction between readers and texts, i.e., the reading process. The availability of sufficient amount of reading materials is a predictor for the success of an extensive reading program. In practice, reading materials available to readers are: print discourse, which includes graded readers, textbooks, newspapers, magazines and other genres, and non-print discourse, which includes the Internet and other electronic media. The advantages and disadvantages of each source are detailed in subsequent sections.

2.3.1 Extensive reading by means of graded readers

What are graded readers?

Because of the characteristics inherent in extensive reading, graded readers are a natural friend of extensive reading, or to be more exact, extensive reading in English, since graded readers have long been the resources available largely to EFL and ESL learners and are mostly without parallel in other languages (Hill & Thomas, 1988a, 1988b), (but this has changed in the last ten years, see for example CIDEB Black Cat not only for the English language, but also for French, German, Spanish and Italian). The fact is that graded readers prosper in step with the boom of extensive reading. The growing acceptance and popularity of extensive reading witnesses the publication of numerous graded reader series by the major international publishers, for example, the Macmillan Readers since 1969, the Heinemann readers and the Oxford readers since 1988, just to mention a few here because of limited space.

Sharing the benefits as well as problems with extensive reading, graded readers have been given different definitions. The shortest definition comes from Wodinsky and Nation (1988), where they call graded readers as books written within a limited vocabulary. Holding the same view, Allan (2009) believes that graded readers provide accessibility by limiting the number of headwords.

Further research shows that lexical control in graded readers is done not only by merely limiting the “number” of headwords but also by “frequency and usefulness” (Hill, 2008). Nation and Wang (1999) supplemented that the text modification in graded readers involves strictly controlling the lexis and restricting the grammatical structures. The evolution in the definition has made it obvious that to achieve accessibility control in graded readers should be extended beyond lexical items to syntax as well. Hill (2008) supplements the latter definition by adding that grammatical control is achieved through “simplicity”. Further to the inherent linguistic features reflected by lexical and syntactical control, John Milne was the first to introduce the notion “information control” since in the first place a graded reader is presented to readers as a piece of complete reading material instead of fragmented concordances for linguistic and discourse study (Hill, 1997a). In fact, a large proportion of graded readers are simplified version of classics. Thus a comprehensive working definition of graded readers encompassing all the distinguishing properties is: Graded readers are texts written by means of lexical, structural, and information control (Hedge, 1985; Hill, 1997a, 2008).

Criticism of graded readers

Susser and Robb (1990) oppose the indivisibility between extensive reading and graded readers, as the implication in many studies that extensive reading should be confined to graded readers has no basis in theory or practice. The design of graded readers is not in harmony with some second language acquisition theory and current themes in language teaching pedagogy. First and foremost, such readers go against the theoretical support of extensive reading, the Input Hypothesis theory which advocates that the dominant mode of language learning is acquisition, through which L2 learners subconsciously “pick up” the language through input, which should be meaningful, comprehensible and not grammatically controlled. However, as discussed above, one characteristic of graded readers is that they strictly control grammar.

Another criticism of graded readers is that graded readers are not authentic comes from applied linguistics. It is assumed that the inauthenticity of graded readers cannot engender an authentic reading experience, and even worse in any

case, simplifying a text may not necessarily render it more easily understood by the L2 readers (Honeyfield, 1977).

The two criticisms cover the three components involved in the reading process: reading material, reader and the interaction between them. Criticisms against graded readers primarily challenge their authenticity and in turn the rationale of their existence as a reading resource. Secondly, by imposing the criticism on graded readers the opponents or critics of graded readers have stated their position that L2 readers need authentic material to guarantee an authentic reading experience. The third claim they make is that simplification in graded readers does not necessarily smooth the interaction between L2 readers and reading material or benefit L2 readers more than authentic reading materials, such as textbooks and newspapers, do.

In response to the first criticism, it is necessary to visit the dichotomy underpinning it: authenticity and simplification. The adherence to authentic material may be out of a practical concern that authentic materials provide good models of English and will be useful in situations where some teachers' own mastery over English is not always secure. Authentic materials play a valuable role in creating an authentic language environment in the classroom, and exposing the students to the real target language community (You & Chen, 2009). It is expected that the unsimplified authentic text keeps the students in touch with the target language even long after the completion of the course (Greenwood, 1988).

However, as Davies (1984) points out, *authentic* is a term coined for language teaching and applied linguistics instead of theoretical linguistics. According to Maxim (2002), authentic texts are "materials written to be read by native speakers of the language rather than materials written only to teach language" (p.20). On this issue, Widdowson (1979) writes "I am not sure that it is meaningful to talk about authentic language as such at all. I think it is probably better to consider authenticity not as a quality residing in instances of language but a quality which is bestowed upon them, created by the response of the receiver" (p.165). A practical interpretation of Widdowson's view by Davies

(1984) is that simplification is used to make information available to an audience other than the one originally intended. If simplification is defined as making a text accessible to the audience, then perhaps any text may be considered a simplification. Is an English translation of a Chinese novel not an authentic English text? Is a plain version of obscure bureaucratic documents written for a larger readership not authentic?

Connecting authenticity and simplification directly, Lautamatti (1978) argues that simplified texts are used as “a ladder towards less simplified and finally authentic texts” (p.98), supporting fully one notion of extensive reading that it prepares readers for authentic reading. For our discussion the importance of Lautamatti’s remark is that simplification and authenticity can be seen as ends of the same continuum, not polar opposites that are absolutely irreconcilable. On the other end of the continuum, simplicity is not equivalent to inauthenticity. A useful distinction is made by Widdowson (1978) between simplified versions and simple accounts. He suggests that simplified versions are derivations of authentic language data by means of alternative textualization, and on the other hand a simple account represents a new type of discourse rather than lexical and syntactic substitution of the original language data.

In practice, there are two major types of graded readers: simplified versions of classics and simple original stories specially written as graded readers by native speakers of English. According to Widdowson’s (1978) distinction, the simple original type is certainly authentic, while simplified versions of classics and modern novels are not. From this point of view, the statement that graded readers are unauthentic is untenable as at least half of graded readers are authentic reading material according to Widdowson’s (1978) criteria. Further, little research has ever been done on to what extent simplification maintains or robs the original, and so graded readers are not at least theoretically or thoroughly inauthentic.

Out of the theoretical discussion of authenticity of graded readers, a more useful question is raised by Irvine (2006) that whether the language of graded readers conforms to the rules of the linguistic system. Empirical investigation has been

made into authenticity of graded readers. Honeyfield (1977) examines the traditional simplification techniques. The result shows that materials produced by these techniques are significantly different from normal English in terms of information distribution, syntax and communication structure, but it is not the linguistic system that suffers from simplification. In addition to the comparison in terms of the whole linguistic system, Allen (2009) compares the lexical chunks in the corpora of graded readers with those in the British National Corpus by means of examining frequency, type and composition. Results show that the frequency of useful chunks in the graded corpora is comparable with the British National Corpus, an authentic corpus, and thus suggest graded readers provide authentic language input. In essence, “The code is not affected, the learners are not presented with a simpler language system but with a restricted sample of the full system (Tommola, 1979, p. 183, cited in Irvine, 2006).

From the readers’ perspective

Regarding the second view that L2 readers need authentic reading material, the discussion in this section seeks to locate what kind of reading material L2 readers should read and should be able to read. The premise guiding this discussion section is the recognition of the differences between native speakers of English and ESL or EFL readers. The L2 reader is not an L1 reader, and in this project, an EFL reader is far from an L2 peer, let alone an L1 one. The situation of an L1 reader and his reading in the native language can by no means be exactly replicated.

L2 proficiency is a prerequisite for the success in L2 reading (Davis & Bistodeau, 1993; Laufer, 1997; Taillefer, 1996). Understandably, reading unsimplified authentic reading materials is considered to be beyond the ability of learners of low-level second language proficiency (Maxim, 2002). Bamford and Day (1997) suggest that simplified materials are an optimal alternative for learners who are not able to cope with unsimplified authentic reading materials.

Research has shown that to ensure successful reading comprehension, a mastery of at least 95% of the running words is necessary and in later research this percentage rose to 98% (Nation, 2001). If unsimplified texts are the only

authorized reading resource for ESL and EFL learners, they may encounter great difficulty in their daily reading since printed texts contain the most unknown words (see Table 2.1 below). For readers at beginning and intermediate levels this problem is even more acute (Carver, 1994; Nation & Deweerdt, 2001).

Table 2. 1 Frequency of rare words in various sources

| | Rare Words per 1,000 |
|---|----------------------|
| I. Printed texts | |
| Abstracts of scientific articles | 128.0 |
| Newspapers | 68.3 |
| Popular magazines | 65.7 |
| Adult books | 52.7 |
| Children's books | 30.9 |
| Preschool books | 16.3 |
| II. Television texts | |
| Prime time adult shows | 22.7 |
| Prime time children's shows | 20.2 |
| Mr. Rogers and Sesame Street | 2.0 |
| III. Adult speech | |
| Expert witness testimony | 28.4 |
| College graduates talk to friends/spouses | 17.3 |

(Graves, 2006, p. 41)

To function in English, users are expected to have a command of no less than 5,000 word families for oral discourse, and a minimum of 8,000 word families for written discourse (Webb & Rodgers, 2009). In addition to the number of word families, various levels of knowledge are involved if learners wish to claim knowing the words. Taken together, the number of word families and the various levels of vocabulary involved amount to a substantial lexical learning task, which is beyond a considerable number of ESL and EFL learners (Schmitt, 2008). The data support the adoption of graded readers in which the frequency of unknown words is much lower than that in unsimplified authentic texts (Broughton, et al., 1978). Further, controlled lexis in graded readers ensures that readers are repeatedly exposed to the high frequency words, and thus the readers' chance of learning the high frequency words or enriching their knowledge of the high frequency words is much greater given that lexical control in graded readers is accomplished by frequency and usefulness (Nation & Wang, 1999).

In addition to the lexis problem, learners of relatively lower English language proficiency are unable to process the syntactical complexities of unsimplified authentic materials (Klapper, 1992). As Pargment (1943) has suggested, reading

material must be graded by gradually increasing levels of difficulty, and accordingly reading materials adapted to cater for the age groups, the interests and the proficiency levels of the learners produced the best results (Coleman, 1930).

Text-learner interaction

It has been shown that under no circumstance should it be assumed that the quality of graded readers suffer because they are controlled in terms of vocabulary, grammar and syntax, but on the other hand, it cannot be taken for granted that simplified language was necessarily more accessible to readers.

Davies (1984) reports an experiment in which both original and simplified versions of the same passage were given to the same group of learners. Learners' performance on a set of subsequent cloze tests shows that they comprehended the simplified version significantly better than the original. Conversely, Blau (1982) examined the effects of shortened and simplified sentences on the comprehension of ESL students in Puerto Rico. Three versions were prepared. The first and simplest version was characterized as containing only short, simple sentences, the second as containing complex sentences with contextual clues and the third as containing complex sentences without clues. Based on students' performance on multiple-choice questions following each passage, the second version, longer sentences with contextual clues, was found to have been consistently better understood. Thus Blau concluded that "the relationship between syntax and readability is not so strong as may have been expected", (p. 527), and for our purpose this result means that simple syntax is not necessarily more easily understood.

As shown above, both positive and negative evidence regarding the relationship between simplification and readability are available, and thus it cannot be concluded with full confidence whether or not simplified text leads to better readability. What happened to the simplified texts in Blau' study was that nearly all sentence connectors and cohesive markers were deleted, resulting in a lack of inter-sentence connection clues and producing a series of short, sharp, apparently unconnected sentences. This is exactly the kind of simplification which has given

graded text a bad name.

Benefits of graded readers

In addition to the theoretical justification of graded readers, there are further features that help make graded readers easier for the learner. Compared with the shorter texts available online and in newspapers and magazines, graded readers have their own strengths.

The first strength is that graded readers are longer than individual articles. The length of graded readers allows sufficient space for background information, with the assistance of which the people who read graded readers may be able to predict the information flow in the book (Martino & Block, 1992). In this way, lengthy text helps the readers of limited language proficiency to overcome some difficulties in reading (Maxim, 2002). For example, readers can make use of the high frequency words and contextual clues to make sense of difficult lexical and syntactic items. For another example, people will be able to make sense of the whole story by referring to the episodes and events in the text. It must be pointed out that graded readers will not cause the readers to panic because of their longer length than individual articles. Compared with the lengthy books intended for native speakers, writing a graded reader is more than simplification of the language. They are not only shortened and simplified, but also abridged and adapted (Greenwood, 1988).

The second strength is that new words in graded readers are glossed. Jacobs (1991) examines the function of vocabulary glossing in second language reading. 116 college students enrolled in a third semester Spanish course were randomly assigned to two groups, with one reading an unglossed Spanish text and the other reading the same text accompanied by English glosses in the margin. After reading the text, participants were required to write as much as they could recall of the text in English. Results showed a significant effect in favor of glossing, based on which it is suggested that glossing has a useful role to play in aiding reading comprehension. This suggestion is accepted by Karp (2002), who suggests that a word is more likely to be acquired if drawn to the readers' attention after a single exposure, although the role of extensive reading in

incidental vocabulary acquisition has been well established.

The third strength comes with the development of new technology. Nowadays a considerable number of graded readers are presented with multi-media resources, such as tapes and CDs. It is found by Hickey (1991) that leisure books accompanied by tapes result in significant increase in the children's motivation to read and higher reading rate than non-taped books. Besides the improved motivation to read, it is indicated that tapes aid comprehension and improve accuracy in pronunciation. Tadman (1980) discussed how graded readers can be better used with cassette. The proposed options include using cassettes for note-taking practices, using cassettes for listening practice and other activities in practicing language skills, which serve as good complements for the development of communicative competence. These possible applications indicate that multi-media resources provide important environmental support for reading graded readers, and furthermore the multi-media resources increase the readers' exposure to the target language in different forms.

To summarize, graded readers present stories that are not only shortened and simplified, but also abridged and adapted. Many graded readers include a glossary of key words, and questions to improve the learners' comprehension. The use of multi-media resources is part of the effort to provide environmental support for reading. Graded readers are therefore easily distinguished from normal books as they are written exclusively for learners of English.

More empirical evidence

In addition to the benefits noted above, a further motivation for the practitioners to use graded readers result from a practical concern that decoding difficult text is a slow, arduous business often requiring recoding into the L1, particularly among lower proficiency readers (Cohen, 1995). Because reading is slow, not much text is read, and there is not sufficient exposure to language input (Parry, 1991). From another perspective, there are various consequences of reading of texts which are too difficult. Students read very difficult materials so they do not want to read, so their reading proficiency does not improve; they remain at a low level of proficiency so they do not want to read. Finally they get locked into this

vicious circle, which keeps them away from reading and improvement (Nuttall, 1982).

In general, the use of easy materials, or graded readers, improves readers' reading comprehension. Takase (2009) emphasizes the effects of very easy materials for beginners, referring to various series of materials from picture books for L1 children to Graded Readers series, with support from the huge success of the Start with Simple Stories Extensive Reading in Japan. Leow (2009) provides a comprehensive review of the research which investigated the effects of simplified written input on L2 readers' comprehension. The research appears to support the use of simplified written texts to promote an improved comprehension of text content in the classroom setting.

At a micro level, graded readers have a positive impact on the individual components of reading. Tanaka and Stapleton (2007) attempted to provide input in English by means of extensive reading in Japanese high school classrooms. In a five-month period, the researchers provided teacher-made materials to the 96 students in their study, and gave them five to ten minutes for in-class reading. Graded readers were available to students who wanted to read more after class. Results showed that those who read graded readers made significant greater gains in terms of reading rate and comprehension than a parallel comparison group that did not read. In Fujigaki's (2009) study, easy books, or graded readers, were used for weak readers in particular. The feedback from the participants confirms that a content-centered approach and level-appropriate reading materials can positively influence student motivation, and that complementary of carefully designed extensive reading courses into the EFL curriculum of secondary schools is of utmost importance for students to build strong reading skills, confidence, and a love for reading. Kitao et al. (1990) reported that 220 graded readers were made available to 300 freshman and sophomore EFL students over a period of four to six weeks. According to analysis of book reports and outcome of the post-program questionnaire, a considerable proportion of the participants have stronger motivation to read in English.

The above empirical studies highlight the following benefits: Graded readers

build reading speed, lexical speed access, reading fluency and increase motivation (Waring, 1997), which point to the irreplaceable role of graded readers in extensive reading (Bell, 2001). For the present study, the pedagogical implication drawn from the research evidence is that extensive reading programs using graded readers are as effective as, if not more effective than, the traditional reading curriculum using short texts for fine-grained linguistic analysis in terms of improving learners' reading proficiency and other language skills, with an additional merit of being more pleasurable and less laborious. It is reasonably assumed that graded readers should resource the extensive reading program to be implemented in the present study. With reference to the specific issues involved in writing and using graded readers, an observation of the graded reader series published by the major international publishers made by the researcher is presented in Appendix G for reference.

2.3.2 Extensive reading through other print texts

Textbooks

Textbooks are the teaching and learning resource most closely related to teachers and students. Brown (2008) believes that textbooks should promote extensive reading, since extensive reading can be bestowed with greater legitimacy and the practical difficulties involved in the implementation of extensive reading can be alleviated in this way. He proposes both direct and indirect approaches of using textbooks to encourage extensive reading. The direct approach focuses on including extracurricular reading materials, mainly graded readers and excerpts from other written materials. The role of textbooks in the direct approach is to present a recommendation list of graded readers which are relevant to the topic of the module and provide reading logs in which students can record their reading. The indirect approach means incorporating reading activities which are more compatible with extensive reading. The approach implemented by Jensen (1986) is a feasible exacerbation of Brown's direct approach. In this approach, authentic materials are combined with textbooks to create theme-based units. These units comprise of in-class compulsory reading and out-of-class supplementary reading as required coursework. The texts too long to be read in class can also be accommodated in the extensive reading session.

The two studies imply that textbooks cannot function as reading materials in an extensive reading program although they are most closely related to teachers and students. The reason is that textbooks themselves cannot be used as reading materials because the word count of a single textbook is very limited. That is not to say, however, textbooks cut all lies with extensive reading. As suggested by Brown and Jensen, the first role of textbooks in an extensive reading program is to present a recommendation list of graded readers which are relevant to the topic of the module and provide reading logs in which the students record their reading. The second role is to serve as a basis for extensive reading activities. These two roles make textbooks an integral component in extensive reading, and pave the way for the complementary of extensive reading into the curriculum.

Unsimplified authentic texts

Authentic texts are pervasive in extensive reading, and at the same time more realistic than any other kind of print discourse, because readers' being able to deal with unsimplified authentic texts is the ultimate goal of many reading programs (Pino-Silva, 2006). You and Chen (2009) advocate that in the field of L2 learning, authentic materials play a valuable role in creating an authentic language environment in classroom, and in exposing students to the real target language community.

Another two additional benefits closely related to authentic text are culture and literature. The impact of culture on reading is emphasized by some researchers (e.g., Day & Bamford, 2000). For example, full-length literary texts have been found helpful in getting students to become independent readers (Ewert, 2009). Fredricks and Sobko (2008) investigate the role of culturally relevant texts in extensive reading. They provide authentic English novels to EFL adult learners in Tajikistan, where extensive reading and other meaning-focused methods are not common. Data collected over eight weeks via students' reflection, instructors' observation and interview showed that the students deemed culturally relevant texts may contain topics previously unexplored in EFL class settings. The inclusion of culturally relevant texts in a course or club may facilitate rich discussions and increase the students' motivation to read. Strong (1996) encouraged extensive reading of multi-cultural literature, as literature is expected

to offer the opportunity to explore cross culture values. Stuart (1990) implemented an extensive reading program with a group of university students in mainland China. In this program, the participants first translated the Chinese folktales into English, and then read simplified novels and finally moved to read unsimplified American and British literature. Approximately fifteen minutes was devoted to students retelling what they read. The final results were highly positive. Reading folktales from their own culture, which many of them are intimately familiar with, brings back a host of pleasant associations and a new appreciation of English.

Despite the numerous benefits, various reading difficulties were reported in studies using unsimplified authentic texts. In Alshamrani's (2003) study, which exposed the students to unsimplified authentic texts in a three-month ESL course called the Reading Club, the reasons or sources of reading difficulties included the existence of new words, technical terms and long sentences, as well as the lack of familiarity with and the background knowledge of the topics. The lexical difficulty that the readers would encounter in reading unsimplified authentic texts was not unique to Alshamrani's study. It is also shown in Chew's (2003) study that highly interesting unsimplified authentic texts as extensive reading materials has a positive effect in improving students' motivation to read, but it works more effectively with a careful plan of implementation and supporting strategies, including direct instruction of common words found across books and identifying the frequency words found in a series of high-interest books by collaborative efforts of the teacher and students.

Lexical difficulty can give rise to a series of problems if the intended audience of an extensive reading program is readers of relatively lower English language proficiency. If so, solutions to the lexical problems such as direct instruction of common words found across books as suggested by Chew (2003) cannot work effectively. If the number of books provided in an extensive reading program is large, the instruction involved will be too great to attain. Similarly, the direct instruction of common words found in one or a small number of books is possible, but it will severely limit the number of books that can be provided to the readers. This dilemma suggests that unsimplified authentic texts have

irreplaceable advantages, but they may not be appropriate for the present study because of the English language proficiency level of the participants.

Newspapers, magazines and other genres

Yu (1995) suggests that many of the principles of extensive reading can be applied to other materials, such as brochures and leaflets in addition to newspapers and magazines. Compared with graded readers and unsimplified authentic texts, newspapers and magazines are more appealing to readers in terms of length, topics, personal interest and cost. The short and pithy articles in newspapers and magazines lower the readers' affective filter towards lengthy texts. The large variety of topics in newspapers and magazines are more likely to cater for the readers' individual reading interest. The extensive sales channels of newspapers and magazines make them more accessible to the readers, and they enjoy natural advantages of being cheaper and more portable.

Positive findings are reported in empirical studies using newspapers and magazines. Constantino (1995) reported from his one-semester reading class for adult that magazine articles helped the students make progress in terms of reading, writing, comprehension and confidence. Barrett et al. (1992) designed a course for high-intermediate to advanced ESL learners, in which students did self-selected reading from newspapers, magazines, books, and even academic journals and textbooks. Using this approach, students became empowered to read materials written for native speakers of English with more ease and confidence.

China Daily is the most widely circulated English language newspaper published in mainland China. Claiming itself to be a guide to official policies, its intended readership is the foreign residents in China and the Chinese citizens with certain degree of English language proficiency. *China Daily* describes its goal as to present news in China to a unique group of readers in an objective manner. *Shanghai Daily* is a newspaper in English published by Wenhui-xinmin United Press Group. It is in principle intended for the overseas investors and tourists to Shanghai with a particular purpose of promoting a positive image of Shanghai, and thus the content of *Shanghai Daily* is exclusively centered on Shanghai and its surrounding region. Primarily, the readership of *China Daily* and *Shanghai*

Daily includes the increasing number of foreigners in China, which implies that high English language proficiency and rich background knowledge about the topics are required to cope with the texts. Using them for an extensive reading program for secondary EFL students would predictably lead to poor results.

Shanghai Students' Post is the first English-language newspaper specifically for the secondary school students and teachers. In order to satisfy the needs of both students and teachers, *Shanghai Students' Post* presents a mixture of reading texts and exercises. Part of this eight-page newspaper is devoted to short articles for free reading, and the topics of the articles cover sports, fashion, entertainment and others that may interest the students. The rest of the newspaper publishes something closely related to classroom teaching, including teaching plans for a specific unit in the textbook, discussion on a grammar structure and corresponding exercises. The proportions of the newspaper devoted to articles and teaching-related content are not fixed, but usually the ratio between the two is roughly equal. The articles in *Shanghai Students' Post* seem to appeal to the secondary school students, as they cover a comprehensive range of topics and they do not cause too many reading difficulties to the students because the articles are specifically written for this age group. The reason for excluding this newspaper from the present study is that its word count is very limited. Given that it is published once a week and several pages are devoted to teaching related content, the situation becomes more acute.

Summary

Extensive reading mainly uses graded readers at the beginning of the reading program but progresses to more genres such as journals and magazines, and even texts written for the peer native speakers (Hill, 1997a). The general trend observed is that reading materials are in line with readers' ages and linguistic competence. Graded readers are common among young readers, while full-length and unsimplified books are suitable for adult readers. The trend is highly consistent with EPER's recommendation that children's fiction and magazines are for primary school students, teenage fictions and magazines for secondary school students and adult fictions and journals for adult students. The materials themselves may not be inappropriate, but need be selected to cater for readers'

age groups, competence and interests.

2.3.3 Extensive reading through the Internet and other electronic media

Reading through print discourse has always been the most common reading activity for EFL learners; however, with the development of information technology reading in the 21st century is no longer confined to print discourse (e.g., Cobb, 2005; Krashen, 2007). The technology is changing the nature of literacy and so the ways of conducting reading has to be changed accordingly (Lee, 2005). Parents and teachers complain that students play too many computer games but read too little, and so it is interesting to see how much the information technology can contribute to extensive reading.

Another reason for the rise of extensive reading through Internet and other electronic media comes from a weakness inherent in extensive reading through print discourse itself. The availability of a large quantity of reading materials is one of the pre-conditions for the success of an extensive reading program. However, in some countries and regions such materials are not easily accessible to readers, or a library does not always have the varieties of books that once exists; and further, the physical obstacles, such as the need to establish and maintain a library of readers hinders extensive reading (Robb, 2009). In such situations the Internet emerges as an optimal alternative for paper reading materials. This availability of large quantities of reading materials, which are accessible from anywhere in the world, makes extensive reading easier and more convenient. This quality of extensive reading through the Internet is of particular value in Mikulecky's (2009) study. The Internet-based children' and young adult literature for extensive reading reduce cost and eliminate adult embarrassment because of the widespread availability of simple literature and children's stories on the Internet.

In addition to the availability and quantities of reading materials, computers can do much more than simply provide and distribute the materials. Author and Stevens (1996) in the 1990's propose that computers be able to not only make available a huge quantity of reading materials but more importantly enhance the learning of the texts in various ways. In the 2000's, a systematic summary of

online reading made by Al-Rajhi (2004) suggests its following advantages: (1) relative accessibility and flexibility in terms of time and space, which have contributed to the boom in e-books and e-journals around the world, (2) the possibility of inclusion of multi-media elements like sound and video clips which cannot be presented in printed texts, (3) opportunities for readers to read and write in real-life contexts, e.g., BBC, CNN and other similar sites provide news in written texts, and after reading these readers are given the chance to comment in writing about what they have read through created channels of interaction, (4) timely and appropriate computer-assisted communication where readers and teachers share what they have read, and (5) accessibility and flexibility of background information of the topic being read in this computer-assisted reading environment. The on-line reading platform introduced and used in Sun's (2003) study adds more advantages to those summarized above. The teacher and students interfaces make possible the interaction between teachers and students during reading. In order to cater for the specific needs of individual readers and offer spontaneous help to students in their reading process, the following reading aids were integrated into the system: concordance help, stage-by-stage reading strategy training, and text annotation functions. Results show that the users hold a positive attitude to this system.

Despite the enumerated benefits of online reading, there are some people who still prefer print texts to electronic texts for different reasons. The first objection is that some people consider reading online to be more challenging, less interesting, and less persuasive and time consuming. The second objection is that extensive reading online is different from the paper-based extensive reading programs in a distinct aspect: there is no teacher pre-selection to secure the linguistic suitability of the reading materials, i.e., whether the learners can read at a level that they are comfortable with, or at their *i-1* level (Arnold, 2009).

The third objection is that benefits gained through online reading are not as solid as expected. In Alshmaikh's (2004) study, eighteen advanced ESL MBA students participated in an eight-week on-line reading program. The participants were divided into two groups as readers and non-readers. In this eight-week program, the participants in the reader group did extensive reading through the

Internet about a comprehensive range of topics, and were asked to keep vocabulary logs and writer journal entries. Both groups completed a questionnaire and took the pre- and post- vocabulary tests, which measured the depth of knowledge of the lexical items they attempted to learn in this program. Findings from the questionnaire and tests indicate that vocabulary acquisition through Internet reading was motivating but did not lead to improved language proficiency.

2.4 Extensive reading schemes

The objective of an extensive reading program is to improve the participants' language proficiency in the target language, with particular concerns on EPER level and speed (Hill, 1997a; Scott-Conley, 2011). According to Yu (1995), the aims of an extensive reading scheme are to help students to acquire the ability to read independently, develop a habit of reading in English, foster a love of reading and improve students' language proficiency through exposure to a wide variety of books. Nuttall (1996) mentioned that a reading program aims to "enable students to enjoy reading in the foreign language, and to read without help unfamiliar authentic texts, at appropriate speed, silently with adequate understanding" (p. 36). Table 2.2 below presents a summary of the key components of an extensive reading program by Scott-Conley (2011). With a clear mapping of the key components of a well-established extensive reading program in mind, extensive reading practitioners are able to make optimal choices even in the planning stage.

Table 2. 2 Components of an established extensive reading program

| | |
|---------------------------|---|
| Library | <ul style="list-style-type: none"> ● Includes a variety of high interest graded readers (and possibly magazines and not EFL materials) organized by level. ● A clear check out system exists. |
| Expectation for students | <ul style="list-style-type: none"> ● Read 500 pages each semester at the appropriate level. ● Demonstrate that they have read. ● Record the number of pages read, keeping a running total |
| Expectations for teachers | <ul style="list-style-type: none"> ● Address student motivation ● Begin ER using a class set ● Teach reading strategies ● Check reports of student reading ● Support students in self-selecting titles ● Do follow up activities in class to expand |
| Assessment | <ul style="list-style-type: none"> ● Pre and Post tests of level, speed, and enjoyment. |
| ER coordinator | <ul style="list-style-type: none"> ● In charge of books, curriculum, and teacher support. |
| Environment | <ul style="list-style-type: none"> ● Posters, cartoons, book reviews, student comments to promote reading. |

(Scott-Conley, 2011)

2.5 Examples of extensive reading schemes in Asia

In order to help students to realize the importance and long-term benefits of extensive reading, extensive reading schemes have been set up worldwide, from the inner circle, for example, England and New Zealand (e.g., Hafiz and Tudor, 1989), to the outer circle, for example, Singapore (e.g., Mok, 1994), and the extending circle, for example Japan (e.g., Takase, 2005, 2007a, 2007b) (see Section 2.2 for more extensive reading schemes). As the present study is going to be carried out on senior one students aged 15 to 16 on average who have just started senior high school, it is useful for comparison to look at studies in which the participants are of the same age and in similar contexts. The following are the relevant research studies of extensive reading of English conducted on primary and secondary students in Asia.

2.5.1 The Hong Kong Extensive Reading Scheme in English (HKERS)

HKERS is the biggest government-funded extensive reading program that has ever been undertaken. It was first introduced into the lower secondary levels early in 1991, and then was progressively extended to senior forms in 1997 (Green, 2005). In four years starting from 1997, it was extended to all secondary levels.

HKERS was established taking guidelines from EPER. The first step of HKERS is to determine the students' entry level to this scheme by administering the EPER Placement Test to the participants. Then a total of 400 books are distributed to eight levels of difficulty with 50 titles at each level in light of lexical difficulty and book length. The class time suggested by the HKERS was two to three periods, which lasted from 80 to 120 minutes per week or cycle. There was no need of a special room for the HKERS lessons. There were subsequent tests to monitor the students' progress, after passing which they were promoted to a higher level; otherwise they would be demoted to a lower level.

Since the implementation of HKERS, a number of studies on extensive reading in English in the Hong Kong context have been conducted (e.g., Green, 2005; Irvine, 2006; Lai, 1993a, 1993b; Tsang, 1996; Yang, 2001). Among them, most

of the studies focus on the language benefits of extensive reading. Lai (1993b) reported significant growth in reading comprehension, reading speed and writing among the 200 secondary school student participants after a four-week summer reading program in which students needed to write simple book reports and engage in class activities, such as reading poems aloud and playing language games, other than reading books in class and at home. More than ten years later, the L2 English writing of students participating in HKERS was compared to that of non-participant students in Irvine's (2006) study. Samples of timed narrative writing from 392 students in Secondary 2 and 3 were rated. The results indicate that extensive reading in an L2 may benefit language development in quite specific aspects, i.e., fluency, complexity and coherence. At low levels of L2 competence, extensive reading may also accelerate the acquisition of basic grammar through frequency effects.

Despite the studies indicating the language benefits of extensive reading, however, HKERS seems to yield a rather disappointing result in helping students to love reading and develop a reading habit. Storey et al. (1997) mentioned that primary and secondary students in Hong Kong have not developed a habit of reading English books. In a large-scale study comprising a total of 1,000 teachers and 10,000 students from both primary and secondary schools using questionnaires to elicit their views on extensive reading and extensive reading scheme, it was found that both primary and secondary school students were not interested in reading English. Primary school students' interests in reading declined substantially from 87.1% in Primary One to 18.9% in Primary Six. A dramatic drop of reading interests was also reported among secondary schools students since the percentage of students who claimed to have the habit of reading English leisure books slipped from 63.3% at Secondary One to 17.6% at Secondary Seven. In another study conducted by Yu (1999) to evaluate the effectiveness of HKERS in five secondary schools, similar results were also noted. Her research findings showed that although 68.5% of the experimental group believed that the scheme had increased their interests in reading, the remaining 30% of students reported that the scheme had not brought any benefits to them within the two years since its implementation and only 27% of the experimental group regarded "Reading in English" as their favorite pastime.

These two studies show a common phenomenon that both primary and secondary students seem to have a low motivation for reading English. One possible reason for such a low motivation for reading as suggested by Storey et al. (1997) is the lack of interesting books offered to students. With such a low motivation for reading and a lack of suitable books, it seems difficult to promote reading among students.

Besides students' low motivation, another great hindrance for HKERS is the problems emerging from actual implementation of HKERS. The actual implementation of HKERS was strikingly varied from the official guideline as summarized in Green (2005). First, in the overwhelm majority of schools extensive reading was not implemented as teacher-guided activities for checking students' progress and helping them locating their level-appropriate books; conversely, a typical scene was that around 40 students sit quietly in their own seat reading one of the 50 titles of their own level throughout the 80 to 120 minutes with teachers sitting in front of the whole class monitoring the students' progress and making sure rules of silence was kept. The outcome of this is usually a book report. The opportunity for sharing the reading among students is excluded by such implementation of HKERS in the classroom.

These problems emerging from the actual implementation of HKERS in the first place indicates that there should be some suggested changes for HKERS. For example, a separate class time should be allocated for teaching students the needed English language reading skills and strategies to take on the challenge of English reading. More importantly, a favorable reading environment should be created for shared reading. To sum up, the HKERS has defects as well as merits. It is definitely a step in the right direction with room for improvement.

2.5.2 Reading and English Acquisition Program (REAP) in Singapore

With an aim to improve the level of English in Singapore primary schools, REAP was first tried in thirty primary school in 1985. Drawing on the successful experience of the Book Flood Project in Fiji (see Section 2.2 for its details), REAP was conducted in several phases in the classroom. In the first phase, i.e., before the pupils were able to read on their own, the Shared Book Approach was

employed. The two core characteristics of the Shared Book Approach were the focus on meaning rather than form and the pure enjoyment of sharing books among peers (Holdaway, 1979). In the second phase, i.e., after the initial reading for enjoyment, pupils were systematically exposed to Enlarged books and guided to a battery of listening, speaking and writing activities, by doing which the pupils practice using English for communicative tasks.

A comprehensive evaluation exercise was administered to measure the effectiveness of REAP. Results showed in 53 out of the 65 comparisons REAP groups significantly outperformed non-REAP groups. Further, pupils who had taken REAP increased their passing rate in the P3 Examination by 10% (Guo, 2007) (see Section 4.4.2 for the detailed discussion on the grammar section in the evaluation exercise).

Several factors contributed to the great success of REAP as summarized in Ng and Sullivan (2001). At the planning stage, the possible problems that teachers may face when implementing REAP were anticipated by the REAP project team members. They worked out two corresponding countermeasures to the potential problems: the workshop for teacher training and practice before implementation and frequent classroom visits by REAP advisors for consolidation of each main technique during implementation. In addition, on-call resource persons were available in the participating schools. The pre-REAP workshop prepared the teachers for both principles and implementation of REAP, most of which were previously known to the teachers as obscure academic theories. The advisory classroom visits assured the teachers that these visits aimed to support them in implementation rather than judge or evaluate their practice.

Eventually, the success of REAP was extended beyond the greater gains made by the REAP participants than the non-participants in the evaluation exercise. From 1991, REAP was no longer implemented as an independent program but was integrated into the national language syllabus and curriculum materials (Mok, 1994). This is so far the first case that extensive reading was integrated into official English language syllabus and curriculum materials.

2.5.3 Extensive reading in Japan

Two decades ago, Extensive Reading was new to many English teachers in Japan. Realizing that the Grammar-Translation Method should not be the only methodology to be adopted, many English teachers in Japan have been searching for new approaches to motivate students at various proficiency levels to study English. In an effort to raise the students' English proficiency, the government advocated the use of the Oral Communication Approach (Mombukagakusho, 2003, cited in Fujimoto-Adamson, 2006). However, criticism of this approach includes a reduction in student literacy skills, vocabulary size, grammar concept, negative backwash on university entrance examination preparation and students' lowered proficiency level (Noriguchi, 2006, cited in Takase, 2007a). Due to the necessity of implementing a new approach to motivate students, extensive reading emerged as the times demand and has attracted attention as a complementary teaching approach recently, evidenced by a surge in the publication of guides instructing teachers how to implement extensive reading program in schools and the abundance of English books and extensive reading materials available in bookstores (Furukawa, 2005).

However, when implementing extensive reading in the Japanese high school setting, there are teething problems. Some of the teachers who were reluctant to implement extensive reading in their lessons explained their major worries: cost, limited class time, no support from colleagues, the different role of teacher, time-consuming work such as book management, students' lack of reading proficiency, and not enough evidence of its effectiveness, etc. To investigate whether these problems actually exist in an extensive reading program, or in other words, whether practitioners and non-practitioners have different perceptions of extensive reading, a questionnaire survey concerning high school teachers' attitudes towards extensive reading was administered to teachers from all over Japan (Takase, 2007b). The research questions investigated include: what concerns practitioners and non-practitioners have about extensive reading and what positive effects practitioners have found for the use of extensive reading.

The comparison between extensive reading practitioners and non-practitioners revealed some interesting data. As illustrated in Table 2.3, the biggest problem

for both non-practitioners and practitioners was cost or lack of materials, identified by 83.3% of the non-practitioners and 60.9% of the practitioners respectively. It is worth noting that this was the only concern shared by the non-practitioners and the practitioners besides difficulty of the different role of the teacher. The divergence of views between the non-practitioners and the practitioners rose on the following issues. For the non-practitioners, the next three biggest concerns were limited class time, no support from colleagues and difficulty of the different role of the teacher identified by 58.3%, 33.3%, and 25.0% of the non-practitioners respectively, whereas 8.7%, 4.3% and 17.4% of the practitioners saw them as major concerns. For the practitioners, the next three biggest concerns were time-consuming teachers' work, reluctant students who were not interested in reading and difficulty of the different role of the teacher, identified by 52.2%, 39.1% and 17.4% of the practitioners respectively, whereas 12.5%, 0.0%, and 25.0% of the non-practitioners took them as major problems.

The data reveals that overall actual difficulties experienced are fewer than expected. In spite of the common concern of cost or lack of materials, data from the practitioners shows that the main difficulty does not lie in implementing an extensive reading program but managing it. There are several tips that teachers can follow to tackle the problems noted. For example, one way to cope with the demand of materials is to join the Japan Extensive Reading Association (JERA), in which the members can not only borrow a variety of books but also share their ideas about implementing and practicing extensive reading (Takase, 2007b). Another possible solution to this problem is to ask a school-based library to handle the material, which accordingly reduce the burden on teachers. Reluctant students can be motivated by providing them with an abundance of easy materials, and by the provision of in-class reading time for Sustained Silent Reading (Takase, 2005). Teachers can be theoretically and practically trained in various workshops to adapt themselves to their new role in an extensive reading program.

Table 2. 3 Concerns in implementing and running an extensive reading program

| Item | ER | ER |
|---|-------------------|---------------|
| | Non-practitioners | Practitioners |
| 1. Cost or lack of materials | 83.3 | 60.9 |
| 2. Limited class time | 58.3 | 8.7 |
| 3. No support from colleagues | 33.3 | 4.3 |
| 4. Difficulty of the different role of the teacher | 25.0 | 17.4 |
| 5. Time-consuming teachers' work (book management included) | 12.5 | 52.2 |
| 6. No time for teachers' own ER | 12.5 | 0.0 |
| 7. Reluctant students who are not interesting in reading | 0.0 | 39.1 |

(Takase, 2007b, p.8)

Concerning the effectiveness of extensive reading, the practitioners reported encouraging benefits that students enjoyed reading, improved their English proficiency, built greater self-confidence and developed a liking for extensive reading. Furthermore, communication between teachers and students increased, and what was worth mentioning was that teachers themselves experienced the joy of reading (see Table 2.4 below).

Table 2. 4 Reading experience and perceived positive effects of an extensive reading program

| Items | Total (%) |
|---|-----------|
| 1. Students enjoyed reading. | 87.0 |
| 2. Students' English proficiency improved. | 39.1 |
| 3. Students became confident in English. | 30.4 |
| 4. Communication between students and teachers increased. | 26.1 |
| 5. Library books were checked out in quantity. | 21.7 |
| 6. Teachers read a lot of books. | 21.7 |

(Takase, 2007b, p.9)

The pedagogical implication drawn from Takase's study is that actual difficulties involved in running an extensive reading program are fewer than expected, but the coordinator or teachers are advised to be well prepared for two potential difficulties: time-consuming teachers' work and reluctant students who are not interesting in reading. Despite the difficulties, it is encouraging that the results suggest that extensive reading in high school setting is a worthwhile project to undertake. It is well worth the teachers' time and effort to incorporate extensive reading into their lessons.

2.6 Conclusion: guidelines for a school-based extensive reading program in Shanghai

The extensive reading programs around the world have provided both successful

and unsuccessful experiences. First of all, extensive reading is not an all-in-one solution to the various teaching and learning problems. To successfully implement an extensive reading program, the context, i.e., the local educational and social-cultural environment, is one of the most influential factors that must be taken into consideration (Shen, 2008). As introduced in Chapter One in the Chinese context the washback effect of the NMET and the traditional Chinese culture shaping teaching and learning in classroom are the most important factors influencing extensive reading. The increasing proportion of reading in the NMET has provided a favorable condition for extensive reading. Thus the possible constraint on extensive reading in Chinese context is the teacher factor. Given that extensive reading is still young and its history in China is even shorter, English teachers in China may not be familiar with or have a deep understanding of what extensive reading is and the benefits it may bring; neither are they trained in terms of the implementation of extensive reading. Another constraint is that stakeholders of the NMET (see Section 1.2 for detailed introduction of this highest-stakes English exam in mainland China; see Cheng, 2008; Li, 1990; Qi, 2005 for its importance in the Chinese educational system), including the administrators, teachers, students and parents, will have a strong sense of insecurity if the teaching activities are not NMET-oriented. Besides the assumed constraints on extensive reading, there are practical constraints on reading, among which lack of instructional time is the biggest.

Second, an adequate supply of reading materials is the prerequisite. Adequate supply does not mean sufficient copies only but also the number of titles. The experience of HKERS has shown that a lack of sufficient titles can be a threat to an extensive reading program. We should also be reminded that an adequate supply of titles and copies is not an insurance of an extensive reading program. Underuse or ineffective circulation of materials would offset the effort to make them available. Setting up an effective circulation system and bring the students and books together is worth our effort.

Third, given that extensive reading is much younger than traditional instructional approaches, it would not be surprising that students do not know what extensive reading is and how to work with it at the very beginning. They need to be taught

and prepared before they begin their extensive reading journey. The experience of HKERS has informed us of the importance of students' preparation for extensive reading. The successful experience of USSR showed that developing students' habits of sustained silent reading is a prelude to an extensive reading program.

Fourth, the role of assessment is to measure the effectiveness of extensive reading. It should be valid, reliable, standardized and applicable to a larger scale, which would promote the development of extensive reading to a larger scale. However, as noted in the literature review one of the challenges facing extensive reading is that tiny increments of growth are hardly inspiration for curriculum designers and language teachers to implement extensive reading programs into their schools or incorporate in the curriculum. We can never rely on a single assessment tool to detect the increments of growth from extensive reading. The REAP researchers adopted 65 comparisons and the Book Flood Project employed an extensive range of language tests. The second challenge faced by extensive reading practitioners is that there is not yet a universal instrument tool to standardize the effectiveness of extensive reading across different countries and regions. The reported benefits so far are established by self-developed exams or small-scale instrument tools. To draw on the successful experience of previous extensive reading schemes, for the validity of the effectiveness of extensive reading, and also the feasibility of extensive reading assessment, two set of tests, one local and one internationally-acknowledged test for extensive reading, the EPER Placement Test, would be employed in this study.

Having decided on the guidelines for extensive reading in this study, a crucial question cannot be avoided: the working definition of extensive reading in this study. For now a definite answer or a working definition of extensive reading cannot be given for this study. As discussed in the literature, extensive reading is defined from different perspectives. So far there has been little empirical study of extensive reading in Chinese secondary schools, which could provide pedagogical implications for this study. The solution to this problem is to establish two experimental groups in this study, as the Book Flood Project in Fiji. As introduced in Section 2.1.1, how we define extensive reading, especially

extensive reading as an input medium and extensive reading as a teaching approach will greatly influence the way we implement extensive reading in the classroom. Thus, in the present study the treatment adopted for the first experimental group investigates the benefits of extensive reading as a medium of voluntary additional input to the normal curriculum. The treatment adopted for the second experimental group will attempt to use extensive reading as a required component complementary to the standard English language curriculum, investigating the use of extensive reading practices instead in comparison to grammar and translation exercises. The first dimension of the study is to examine the effect of extensive reading by comparing the performance of the two reading groups with one non-reading group. The second dimension of this study is to examine the difference between the two experimental groups, in order to gain insight into the effect and limitations of extensive reading in this context, draw pedagogical implications for subsequent extensive reading research, and provide evidence for the discussion on the less-agreed and discussed features of extensive reading.

3 Design and Implementation of a School-based Extensive Reading Program in Shanghai

3.1 Overview of the program

If an argument is to be made for the effectiveness of extensive reading, it must be well researched in the type of environment in which it will actually be implemented. Therefore, the main focus of Chapter 1 was to explore the context in which the present study is going to be conducted, and the main focus of Chapter 2 was to establish theoretical support for extensive reading and draw pedagogical implications from previous extensive reading studies in order to form guidelines for the present study. Based on the argument made in the previous two chapters, an extensive reading program is going to be implemented in a public senior high school in Shanghai as will be detailed in this chapter. The primary concern of this research undertaking is to examine the significance of extensive reading on senior high school students' English language proficiency and its role in improving English language teaching quality,

As introduced in Chapter 2, how we define extensive reading, especially extensive reading as an input medium and extensive reading as a teaching approach will greatly influence the way we implement extensive reading in the classroom. (see Section 2.1.1 for detailed discussion). Thus, in the present study the treatment adopted for the first experimental group investigates the benefits of extensive reading as a medium of additional voluntary input to the normal English language curriculum. The treatment adopted for the second experimental group will attempt to use extensive reading as a required component complementary to the standard English language curriculum. The secondary research concern involved investigating how extensive reading can be effectively implemented in this research context, and was examined through the use of two different extensive reading methodologies for the two different experimental groups. Outcomes of the present study will be used to recommend feasible changes or modifications in the current practice in English language teaching in Chinese secondary schools.

3.2 Sampling procedures

One benefit of integrating extensive reading into normal curriculum is that, although there is a considerable body of experimental evidence showing that extensive reading does bring benefits to learners' second language proficiency gains, the comprehensive range of contextual constraints on English language teaching in the Chinese context introduced in Chapter 1, including the uniformed curriculum standard, the washback effect of the NMET, limited instructional time and rigid curriculum. It is certainly appropriate to integrate extensive reading into the curriculum (Macalister, 2007). As Asraf and Ahmad (2003) remark, it would be hardly possible for students to give priority to reading over other activities unless extensive reading is integrated into the standard curriculum.

To examine the effect of extensive reading in secondary school, the first step is to establish comparable research groups by strict sampling procedures. After being admitted into the sample school through their performance in the Shanghai Senior High School Matriculation Test, the newly-admitted Senior One students were required take another set of tests before they began their study in the sample school. To make the classes parallel, the sample school has a long tradition of stratifying new students by pre-term tests. Every student admitted to this school is tested on math, English, physics and chemistry. Then they are stratified into eight classes randomly by computer according to the following major criteria: their total scores, their scores on individual subjects tested in the stratification test and their genders. The result of stratification is to ensure that the new classes are parallel in terms of total scores, individual subjects and genders. The pre-tests, which are also a stratification test, were developed by the head of the teaching staff of the school district. The tests were made available to all senior high schools in this school district. Whether schools chose to use the tests or not was optional, but they had no input into the development of the tests.

After the pre-tests and the stratification, the thirty-three students in the researcher's class, i.e., Class 5, were recruited for the study as the participants in one of the experimental groups, i.e., the Complementary Reading group. After the Complementary Reading group was established, a call for volunteers to participate in the present study was circulated to the students in the seven other

classes via their class English teachers. The researcher briefed the class English teachers on extensive reading, and the class English teachers briefly introduced to their students what extensive reading was, the benefits that extensive reading had to offer, how the program would run and what they would be expected to do in this program. After the introduction, sixty-seven students showed their interest and registered their name via their class English teachers. At this stage, the newly-admitted 297 freshmen were in three groups: the class in the charge of the researcher (N=33), which was designated as the Complementary Reading group, the sixty-seven students who volunteered to participate in the present study from the seven other classes, and the remaining 197 students who showed no interest in or willingness to participate in the present study.

The academic officer in the sample school who was in charge of the stratification every year was responsible for establishing the second experimental group, i.e., the Free Reading group and a comparable Control group. In reference to the Complementary Reading group's beginning English language proficiency, namely their performance on the English test in the pre-tests, and gender composition, thirty-three were chosen from the sixty-seven and thirty-three were chosen from the 197 by stratification software. Conditions permitting, an exact match was preferred, namely, the composition of the Free Reading group and the Control Group were exactly the same in terms of numbers of students from different classes as illustrated in Table 3.1 below. The result was conveyed to the candidates via their class English teachers.

Table 3. 1 Composition of the Free Reading group and the Control Group

| No of participants from | The Free Reading group | The Control Group |
|-------------------------|------------------------|-------------------|
| Class 1 | 3 | 3 |
| Class 2 | 5 | 5 |
| Class 3 | 6 | 6 |
| Class 4 | 5 | 5 |
| Class 6 | 3 | 3 |
| Class 7 | 7 | 7 |
| Class 8 | 4 | 4 |
| Total | 33 | 33 |

Note: Class 5 was the Complementary Reading group.

After the Free Reading group was established, the first orientation was given to the tentative free reading participants before the graded readers, i.e., the Oxford Bookworm graded readers, for this program arrived (for detailed information of

the graded readers, please see Section 3.4). This orientation aimed to help the participants appreciate the power of extensive reading, introduce them to the books and program, and assist them in developing their reading efficacy by establishing a daily reading habit, reading at a quick, steady pace, and avoiding dictionary use.

The thirty-three free reading participants gathered in a spare lecture room, listening to a lecture delivered by the researcher in Chinese on what extensive reading was, the benefits that extensive reading had to offer, the necessary extensive reading skills and strategies, how the program would run and what they would be expected to do in this program. After the first orientation, the consent form was distributed to each free reading participant, and they were asked to reconsider their willingness to join and discuss it with their guardians. As the present study would last the whole academic year, it was necessary to confirm that the free reading participants join this program with a genuine and strong interest instead of being prompted by a sudden impulse. This was done to avoid the possible loss of subjects.

The second orientation was given to the free reading participants after the books arrived. The 235 books for the free reading group were stored and displayed in a spare classroom, which was used as the program's library for the free reading group in the present study. The books were introduced to the thirty-three free reading participants in detail, including the levels of difficulty, topics and genres. Afterwards, the details of how the program would run and how the books would be circulated were introduced. The tentative free reading participants' willingness to participate was reconfirmed after the second orientation. All of the thirty-three participants made it clear that they wished to be a participant in the present study, and signed the consent form in the presence of the researcher.

To summarize, the present study was designed with two experimental groups (N=66) and one control group (N=33). The participants in the two experimental groups were selected by means of convenience sampling. The reason for selecting the students in the researcher's class was that they were easiest to recruit for the present study and they were easily accessible to the researcher. The reason for

selecting the participants in the Free Reading group was that they volunteered to participate, so it did not violate the research ethics.

The strongest criticism on convenience sampling is sampling bias, i.e., the experimental population formed by convenience sampling may not represent the entire population, which will result in the limitation in generalizing the research findings and making inference about the entire population based on the findings (Nunan & Bailey, 2009). The limitation in generalization and inference making about the entire population will in turn threaten the external validity of a research undertaking (Cohen & Manion, 1994).

In response to the possible criticism on sampling in the present study, it can be argued that the participants chosen for this study were typical of Shanghai senior high school students. The school in which the present study was conducted was typical of many in Shanghai. It is a public senior high school run and managed by the Board of Education in Shanghai. It has approximately 250 to 350 students in each grade with about 35 students in each class. All of its English teachers graduated from two local teacher training colleges, both of which are reputable in this area in mainland China. They all follow the English Curriculum Standards for Senior High Schools (MOE, 2003), which is universal in Shanghai.

Another reason for locating this school is that it ranks neither high nor low among the senior high schools in Shanghai but is ranked in the middle position, as evidenced by its NMET(s) means and those for all of Shanghai (see Table 3.2 below). Thus the findings of this study may be applicable to a significant number of senior high schools since the students involved in this research are representative of senior high school students in Shanghai in general.

Table 3. 2 Comparison of means of the NMET(s)

| Year | Means of school | Means of Shanghai | Total |
|------|-----------------|-------------------|-------|
| 2007 | 95 | 94 | 150 |
| 2008 | 97 | 97 | 150 |
| 2009 | 107 | 99 | 150 |

Note: Scores in this table are internal data provided by the sample school and the NMETs committee. For confidentiality reasons, only rough scores are presented here.

In addition, the participants chosen for this study were typical of Shanghai senior

high school students. They were all native speakers of Chinese, with Standard Mandarin as the official language and Shanghainese their dialect. All of them were born between 1993 and 1995, and had received formal education in mainland China for nine years in primary and junior middle school following the same syllabus and curriculum standard when the program began. All of them were admitted to the school through their academic performance in the Shanghai Senior High School Matriculation Test. Furthermore, they came from families with similar socioeconomic backgrounds: Most were single children whose parents were factory workers; their family income was average for the city, which should mean their respective families' investment in their education was about the same; and none had studied or traveled abroad or had relatives or friends abroad. Such homogeneity makes for considerable predictability in terms of students' exposure to English.

Another potential variable that may impact students' exposure to English is the teacher design of the present study. The researcher was in charge of the Complementary Reading group only, and the participants in the Free Reading group and Control group were in the charge of the six other English teachers in the same grade. When the EFL context makes for considerable predictability in terms of students' exposure to English out of class, different teachers would lead to different amount of exposure to English mainly because of teaching method in classroom. As introduced in Chapter 1, the Grammar-Translation Method is the dominant English language teaching method in the Chinese context, which is also the case for the present study. Using this method, English is taught in Chinese. In this sense, the Grammar-Translation Method also promises considerable predictability in terms of students' exposure to English.

3.3 Instruments

The tests play a crucial role in the present study. Three precautionary steps were taken to avoid the possible threats to the reliability and validity of the tests. In the first step, measurements that were appropriate for the present study and that had been used successfully in other studies were looked for. In the next step, some of the instruments were piloted to check whether they fit the context of the present study and were able to gather appropriate information to answer the research

questions. Results of the pilot study of the EPER Placement Test are reported in Section 3.3.4. Finally, the reliability statistics of each test were calculated and presented in subsequent sections.

The locally developed standardized tests were employed in this study. Another test, the Edinburgh Project on Extensive Reading (EPER) Placement Test, was also adopted, to provide an additional measure of the outcomes and bring an international dimension to this study. The reason for the absence of TOEFL was that it may not be appropriate for the present study. TOEFL evaluates the ability of the students who wish to study at American universities. The participants in the present study are high school freshmen, and they are going to pursue further education in Chinese universities by NMET three years later. In this situation, IELTS seems a more logical choice, as it is an international standardized test of English language proficiency. The reason for excluding IELTS is the limited research funding. The test fee of IELTS for test takers in mainland China is 1,500 yuan each time. The total cost for the ninety-nine participants taking IELTS twice as pre- and post-tests will be 297,000 yuan, which is equal to about 48,000 US dollars or 30,000 pounds. It is a sum of money that the present study cannot afford. The absence of the internationally-accepted tests, i.e., TOEFL and IELTS, does not threaten the reliability and validity of the present study. The locally developed standardized tests are reliable, as evidenced by the reliability statistics presented in subsequent sections. The EPER test serves as an internationally accepted test. It has been employed in large-scale government-funded extensive reading schemes such as the Hong Kong Extensive Reading Scheme (HKERS). Both its reliability and validity have been internationally acknowledged (e.g., Ambatchew, 2003; Shillaw, 1999).

The first battery of tests employed in the present study comprised of four tests locally developed by the school district where this extensive reading program was implemented. The four tests served as the pre-, mid-, post- and delayed post-tests in the present study to collect quantitative data, which described the academic achievement of the participants in this study and their continued academic development one year after the program ended.

3.3.1 The pre-test

After being admitted into the sample school through their performance in the Shanghai Senior High School Matriculation Test, the new students were required take another set of tests before they began their study in the sample school. To make the classes parallel, the sample school has a long tradition of stratifying new students by pre-term tests. Every student admitted to this school is tested on math, English, physics and chemistry. Then they are stratified into different classes randomly by computer according to the following major criteria: their total scores, their scores on individual subjects and their genders. The result of stratification is to ensure that the new classes are parallel in terms of total scores, individual subjects and genders. The pre-tests, which are also a stratification test, were developed by the head of the teaching staff of the school district. The tests were made available to all senior high schools in this school district. Whether schools chose to use the tests or not was optional, but they had no input into the development of the tests.

In the 2010/11 academic year, the sample school took in 297 new students, who were stratified into eight classes. As planned, the pre-tests on mathematics, English, physics and chemistry were administered in August 2010. The English test was composed of six sections of 100 points in total: (1) grammar (10 points), (2) cloze (30 points), (3) reading in the form of multiple choice (22 points), (4) reading in the form of question and answer (8 points), (5) translation (10 points) and (6) essay writing (20 points). Listening was absent from the pre-test. The pre-test is attached in Appendix A.

In the grammar section, there were ten incomplete sentences for students to complete by choosing the best answer from the four choices provided. In the Cloze Section, students were required to fill in thirty blanks in two passages, and for each blank there were four answers for them to choose from. In the first Reading Comprehension Section, there were three reading passages, and the students needed to answer eleven questions in the form of multiple choice based on the passages, three to four questions for each passage. In the second Reading Comprehension Section, students were required to answer four questions based on one passage. Unlike the previous reading comprehension section in the form

of multiple choices, this section required the students to find the answers in the passage and answer the questions using their own words. Grammatical and spelling mistakes would result in loss of marks. In the Translation Section, the students were required to translate three Chinese sentences into English, with three points for the first two sentences and four points for the third sentence. Grammatical and spelling mistakes would also result in loss of marks. In the Writing Section, the students were required to write a composition of 100 to 120 words on a given topic.

The pre-test was marked by six teaching staff of the sample school; each staff member was in charge of a separate section. Every section was marked once. The reliability statistics of the pre-test presented in Table 3.3 below was calculated based on the test result of the sample school instead of the whole school district. Each school worked on its own to mark the test papers because the use of the tests was optional; because of this the scores of other schools were not available for use in the present study. The markers recorded the score of each subsection instead of the scores on individual items to calculate the total score of each student. In light of this, the reliability statistics of the pre-test was calculated on the basis of the scores of each subsection and the total scores instead of the scores on individual items. There were six subsections in the pre-test. Thus as shown in Table 3.3 below the number of items involved in calculating the reliability of the whole pre-test was seven, i.e., six subsections plus one total score.

Based on these seven items, the reliability of the pre-test was calculated by Cronbach's Alpha Test using SPSS 17.0. The result was that the overall reliability of the pre-test was 0.76. Given that the reliability of the National Matriculation English Test in 2010 was 0.947 (Han, 2010) and the reliability of the College English Test Band 4 (CET 4) and Band 6 (CET 6) is around 0.90 (Zhang, 2010), the reliability of the pre-test is moderately high. Conventionally, if the reliability of a test is higher than 0.70, the test result is valid for comparison between groups; if the reliability of a test is higher than 0.90, the test result is valid for comparison between individuals (Zhang, 2010). In the present study, the reliability of the pre-test was 0.76, which met the requirement for comparison

between groups.

Table 3. 3 Reliability statistics of the pre-test

| Reliability Statistics | | |
|------------------------|--|------------|
| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
| 0.72 | 0.76 | 7 |

3.3.2 The mid-test

A similar test, the first final term test, served as the mid-test at the end of the first academic term, i.e., halfway through the present extensive reading program. It was a large-scale standardized test developed by the Test Bureau of the school district. In Shanghai it is a tradition that every school district develops and administers a large-scale standardized test at the end of each academic term to monitor every school's academic performance. Every school in this school district, public or private, is required to take the test.

The mid-test was administered in mid-January, 2011. It was divided into the following sections: (1) listening comprehension (20 points), (2) grammar (20 points), (3) vocabulary (9 points), (4) cloze (10 points), (5) reading comprehension (16 points), (6) translation (15 points) and (8) essay writing (10 points). The mid-test is attached in Appendix B. About 2000 senior one students in this school district took this test. To ensure the quality of marking, individual schools did not work on their own to mark the test papers. Instead, the teachers from different schools in this school district cooperated to mark the test papers under the supervision of the head of the teaching staff. The marking procedures were as follows.

1. After the test ended, the test papers were collected by the officer of academic affairs in each school, and then were sent to the test center of the school district. A technical team scanned all the papers collected from every school to the central computer.
2. The multiple-choice parts were marked by the marking machine twice.
3. To mark the translation and writing sections, on a certain date all the senior high school English teachers in this school district gathered to mark the papers. The teachers were divided into several groups randomly. Each group was in charge of one part. For example, two teachers were in charge of

Sentence One in the translation part. Apart from the teachers in charge of translation, other teachers were in charge of the compositions.

4. Every teacher sat in front of their own computers, and the central computer randomly distributed test papers to them. Only the part to be marked was visible to the markers. The compositions were presented to the markers in charge of compositions only. For example, the composition markers had access to their own parts, compositions, only. The other parts of the papers, translation and the names of the students and which schools they came from, were invisible to them. They were unable to choose which paper to mark, and could not see how the students did in other parts, either.
5. Having read the students' compositions and made a decision on the marks, the markers inserted the marks into the computer. The central computer recorded the marks that every student got, and calculated their total marks automatically.
6. To ensure a reasonable rater reliability of marking translation and composition, the head of the teaching staff supervised the whole marking process and made a series of guidelines. In marking the translation section, the primary principle was that a grammar mistake costs one point and a spelling mistake half a point. Other minor issues that emerged in the course of marking the paper were discussed among makers and agreement was made before the final decisions on the marks were made. For example, it was noted that some students did not use the key words or phrases given in the brackets. The decision was one point was deducted because of the no use of the given words and phrases even if the translation was completely grammatically and lexically correct. A further action taken to ensure the reliability of marking translation was that each sentence was marked twice by two different markers. An error rate of one point was set by the head and was agreed by all the markers. If the first marker rated a sentence 2 and the second marker scored it 3, the error was one. In this situation the error was within the acceptable error rate, and thus no further action would be taken. Accordingly, the students gained 2.5 points for this sentence, the average of the two marks. If the two makers disagreed with each other on a sentence, and their disagreement exceeded one point, the controversial sentence would be presented to a third marker. In this situation, the final mark of the

sentence would be the average of the three.

The steps taken to ensure the reliability of marking compositions were approximately the same as those in marking translations. In this calibration session, the first step was to randomly select three samples from the test papers to be marked by the head of the teaching staff. The three samples were photocopied and distributed to each marker. All the markers in the first place read the three samples and made their own judgment. Afterwards, they reported their marks and discussed with each other to work out common standards for marking. After the common standards were worked out, the next step taken by the head of the teaching staff to ensure the reliability of marking was to set an error rate. Given that the total mark of a composition was ten, the acceptable error was set as two points. One composition must be marked by two markers. If the first marker rated it 5 and the second marker scored it 6, the error was one. The error was within the acceptable error rate, and no further action was needed. Accordingly the student gained 5.5 points for the composition, the average of the two marks. The situation described above was 97% of the cases. In the remaining 3% of the cases, the two markers disagreed with each other, and their disagreement exceeded two points. In this situation, the controversial composition would be presented to the third marker. The third marker rated it for the third time. He or she could agree with either side, or made his or her own decision. If the third decision was within the acceptable error rate with the previous two, the points for the composition would be the average of the three. There was no case in which the three markers disagreed with each other.

The reliability statistics of the mid-test presented in Table 3.4 below was calculated based on the test results of the sample school. The number of items involved in calculating the reliability of the whole mid-test was eight, i.e., seven subsections plus one total score. Based on these eight items, the reliability of the whole test was 0.82.

Table 3. 4 Reliability statistics of the mid-test

| Reliability Statistics | | |
|------------------------|--|------------|
| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
| 0.75 | 0.82 | 8 |

3.3.3 The post-test

The second final-term test functioned as the post-test at the end of the academic year, also the end of this program. It was identical with the mid-test both in form (see Appendix C) and marking procedures⁸². The reliability statistics of the post-test presented in Table 3.5 below was calculated based on the test results of the whole school district with 1979 test takers. Unlike that in the pre- and mid-test, the number of items involved in calculating the reliability of the whole post-test was eighty-one, i.e., seventy-five multiple choices, five translations plus one composition. Based on the eighty-one items, the reliability of the post-test was 0.90.

Table 3. 5 Reliability statistics of the post-test

| Reliability Statistics | | |
|------------------------|--|------------|
| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
| 0.90 | 0.90 | 81 |

3.3.4 The delayed post-test

The final-term test at the end of the second academic year, which was administered one year after the program ended, served as the delayed post-test. Similar to the mid- and post-tests, the delayed post-test was a large-scale standardized test developed by the Test Bureau of the Shanghai school district⁸³. It was similar to the mid- and post-tests both in form and marking procedures. It was divided into the following seven sections: (1) listening comprehension (30 points), (2) grammar (16 points), (3) vocabulary (9 points), (4) cloze (15 points), (5) reading comprehension (35 points), (6) translation (20 points) and (7) essay writing (25 points).

A notable difference between the delayed post-test and the mid- and post-tests was that the total score of the delayed test was 150 while that of the mid- and post-tests was 100. The explanation for this difference is that the delayed post-test was identical with the NMET (Shanghai Version) (see Section 1.4 for a more detailed discussion of the NMET (Shanghai Version). It is a common

⁸² The essays in the post-test were rated by the same group of markers in the mid-test.

⁸³ The delayed post-test is not attached in the appendix, because the researcher is not authorized to release the test paper.

practice that in Shanghai the large-scale standardized English tests strictly follow the NMET (Shanghai Version) from the second academic year, which confirms the prominence of the NMET in mainland China from a different perspective.

There was also a difference between the delayed post-test and the mid- and post-tests in terms of marking procedures⁸⁴. When marking the essays in the mid- and post-tests, the markers gave a total score for the whole piece of writing without giving further information on sub-components of the essay, such as language quality and organization. In the delayed post-test, the total score of the writing section was twenty-five points, which allows for a more finely-calibrated grading system than the holistic ten-point system used in the mid- and post-tests. The markers were requested to rate the essays in terms of content, language quality and organization. This meant that the markers gave three sets of scores when rating the essays: ten points for content, ten points for language quality and five points for text organization. The three scores yielded a total possible score of twenty-five points.

3.3.5 The EPER Placement Test

Another test selected for collecting quantitative data was a standardized international test developed by the Edinburgh Project on Extensive Reading (EPER) of the Institute of Applied Language Studies of the University of Edinburgh⁸⁵. As specified in the booklet of EPER Placement Test, the EPER Placement Test is used to indicate a student's entry level into an extensive reading program.

The eight levels set by EPER are Levels G, F, E, D, C, B, A and X, from low to high respectively. These eight EPER levels can be matched with either vocabulary levels, student levels or exam levels as shown in Table 3.6 below (for more detailed matching information, please refer to Hill, 2008). For vocabulary levels, Table 3.6 suggests that a student who reaches a particular EPER level

⁸⁴ The essays in the delayed post-test were rated by the same group of markers in the mid- and post-tests.

⁸⁵ The EPER Placement Test B is not attached in the appendix, because the researcher is not authorized to release the test paper.

should be able to read fictions written by the corresponding vocabulary level fluently, not the capacity to read with the help of a dictionary. For student levels, Table 3.6 suggests that the overall English language proficiency of a student who reaches a particular EPER level should be approximately equivalent to a matched student level, for example, elementary or intermediate. For exam levels, Table 3.6 shows that a student who reaches a particular EPER level should be able to obtain the corresponding score in those exams. In the present study, we match the participants' EPER levels to vocabulary level, i.e., at which vocabulary level the participants should be able to read comfortably, for the purpose of locating appropriate reading materials for the participants in the present study. In subsequent presentation and interpretation of the EPER test results, students' EPER levels are interpreted as their ability to read comfortably at a certain vocabulary level (see Section 4.4 for detailed discussion). One more note to make here is that the number of words for the EPER levels refers to headwords (see Section 3.4 for the discussion of the concept of headword).

Table 3. 6 Matching EPER levels with vocabulary levels, student levels and exam levels

| EPER Levels | Average vocabulary | Student level | IELTS | TOEFL | TOEIC |
|-------------|--------------------|------------------|-------|-------|-------|
| G | 300 | Elementary I | / | / | / |
| F | 500 | Elementary II | / | 350 | 150 |
| E | 800 | Intermediate I | / | 400 | 300 |
| D | 1,200 | Intermediate II | 4.5 | 450 | 450 |
| C | 1,600 | Intermediate III | 5.0 | 480 | 530 |
| B | 1,900 | Upper I | 5.5 | 520 | 650 |
| A | 2,200 | Upper II | 6.0 | 550 | 730 |
| X | 3,000 | Upper III | 6..5 | 580 | / |

(adapted from Hill, 2008, p.199)

In terms of test format, the EPER test is a cloze test where a series of twelve reading passages are arranged in order of increasing difficulty. Each passage has from ten to thirteen words taken away and the test takers are required to read the passage and restore the missing words. For each of the blank there could be more than one correct answer. Such a task is largely bottom up and emphasizes careful passive decoding and local comprehension at the micro linguistic level (Ambatchew, 2003).

The total number of items is 144 and the duration of the test is 60 minutes. On the first page, clear instructions about the test are provided. There is also an

example of a passage with five words missing and then the answers are given. A wrong answer is also explained because the test taker is required to fill in the blanks with only one word and the wrong answer is composed of two words. Even if it makes sense it is wrong.

For Example,

Passage

“What are the two dogs doing in road?” said Mary.

“They fighting,” replied David. “Can hear them, Mary?”

“Yes, I can them easily but I want to them also.”

Answers:

| | |
|---|------|
| 1 | the |
| 2 | are |
| 3 | you |
| 4 | hear |
| 5 | see |

In ‘look at’ is wrong because you must fill each blank with only ONE word.

There are five versions of the EPER Placement Tests. Test B was employed in this study because it is the most reliable among all the versions (D. Hill, personal communication, April, 2009). The only possible shortcoming could be that the test and its content may not fit with the Chinese context, so it had to be piloted and validated. To ensure that this test was compatible with the Chinese context, the EPER Placement Test B was piloted with 321 senior one students in the sample school in December 2009. The results are presented in Table 3.7 below:

Table 3. 7 Results of piloted EPER Placement Test B in 2009

| EPER level | A | B | C | D | E | F | G | Total |
|------------|------|------|------|------|-----|-----|-----|-------|
| Headword | 2200 | 1900 | 1600 | 1200 | 800 | 500 | 300 | / |
| No. | 7 | 89 | 148 | 61 | 10 | 4 | 2 | 321 |

As can be seen from the results, the EPER Placement Test is capable of drawing out the range of reading abilities of Chinese students. The reading ability of the students ranged from reading at 3000 headwords level to reading at 500

headwords level. The scale is capable of successfully distinguishing the EPER levels of the various students. Consequently, the test was considered suitable to apply to the Chinese context.

3.4 Reading materials

After the groups had been established, the next step was to locate appropriate reading materials for the participants. Selection of extensive reading materials was based on the following two criteria: (1) readers' linguistic competence, and (2) readers' needs and interest, namely the suitability of content.

Regarding the readers' linguistic competence, the EPER Placement Test B was administered in December, 2009 in the sample school to the senior one students as a pilot study (see Section 3.3 for introduction of this pilot study), and re-administered in September 2010 to all the newly-admitted senior one students to investigate the participants' linguistic competence and their entry level to this program. According to the test results presented in Table 3.8 below, the vast majority of the students, i.e., 95.01% in 2009 and 93.54% in 2010, in the sample school were able to read at the 1200 headwords level.

Headwords noted above refer to word families. In vocabulary studies, articles, propositions, pronouns, conjunctions, auxiliaries, etc, are often referred to as function words, whereas nouns, verbs, adjectives and adverbs are known as content words (Read, 2000). Function words are also referred to as closed word classes, because these word classes can hardly have new members. For example, new articles, proposition, pronouns, conjunctions, auxiliaries have not been coined, although some of them change their spellings over long periods of time. In comparison, content words are called open word classes, because new words have been created with the development of science and technology, such as *corpus* and *Internet*. In addition to its openness to new members, another characteristic of content words is that content words appear in a variety of forms. For example, there is *laugh*, but there are also *laughs*, *laughed*, and *laughing*; similarly, there are *happy*, and also *happiness* and *happily*. The two examples show that base words may not only take inflectional endings but also have a variety of derived forms. Despite the inflectional endings and deviations, such a

set of words share an underlying meaning, and thus known as a word family (Bauer & Nation, 1993). The Oxford Bookworm employed the concept “headword” in this series. Readers are supposed to know *laugh*, as well as *laughs*, *laughed* and *laughing*; similarly, readers are supposed to know *happy*, as well as *happiness* and *happily*. This concept also applies to the EPER Placement Test.

Table 3. 8 EPER levels of Senior One students in the sample school in 2009/10

| EPER Level | | X | A | B | C | D | E | F | G | Total |
|------------|--------|------|------|-------|-------|-------|------|------|------|--------|
| Headword | | 3000 | 2200 | 1900 | 1600 | 1200 | 800 | 500 | 300 | / |
| 2009 | Number | / | 7 | 89 | 148 | 61 | 10 | 4 | 2 | 321 |
| | % | / | 2.18 | 27.73 | 46.11 | 19.00 | 3.12 | 1.25 | 0.62 | 100.00 |
| 2010 | Number | 3 | 15 | 69 | 121 | 67 | 16 | 3 | / | 294 |
| | % | 1.02 | 5.10 | 23.47 | 41.16 | 22.79 | 5.44 | 1.02 | / | 100.00 |

Regarding the suitability of content of the reading material, the researcher read through all the major International series available in the EPER database, including the Black Cat series, the Cambridge series, the Heinemann series, the Longman series, the Macmillan series, the Oxford series and the Penguin series, to locate appropriate graded readers for the participants in the present study. It was decided that the Oxford Bookworm (OBW) series would be used in the present study after the researcher’s reading across different series.

There are three reasons why the OBW was chosen. The first reason is that the OBW series is praised as the most consistent of all series in terms of language control, length, and quality of storytelling by Hill (2008) compared with other series which have titles of very uneven quality. The second reason is that the OBW series is familiar to the students in mainland China since it is one of the earliest graded reader series available in mainland China. Thirdly, the Oxford University Press had a branch in Hong Kong. The books were easily transportable to mainland China from Hong Kong.

The OBW series has 235 titles which are distributed to seven levels of difficulty: Starter (Level 0) (250 headwords), Level 1 (400 headwords), Level 2 (700 headwords), Level 3 (1000 headwords), Level 4 (1400 headwords), Level 5 (1800 headwords) and Level 6 (2500 headwords). The distribution of the 235 titles to the seven levels is approximately equal, with 25 titles at Starter, 41 titles at Level 1, 47 titles at Level 2, 43 titles at Level 3, 39 titles at Level 4, 21 titles at

Level 5 and 19 titles at Level 6. According to the pre-EPER test result, the vast majority of the senior-one students in the sample school are able to read comfortably at 1200 headword level, corresponding to Level 2 and Level 3 in the OBW series. In terms of the readers' needs and interest, the 235 titles in the OBW series are distributed to nine genres, with 36 *Classics*, 24 *Crime & Mystery*, 24 *Factfile*, 31 *Fantasy and Horror*, 44 *Human Interest*, 10 *Playscripts*, 36 *Thriller and Adventure*, 18 *True Stories* and 12 *World Stories*. This variety of genres and topics were expected to cater for students' diversified needs. Eventually, all the 235 titles in this series were read through by the researcher to ensure that they were linguistic and content appropriate for the participants in the present study.

3.5 Control of extraneous variables

Before the new term began, it was arranged that the researcher would be appointed as the English teacher for Class 5. Six other English teachers would teach the other seven classes, and one of them would teach two classes and the other five would teach one class each. Class 5 was then designated the Complementary Reading group, and students selected from the other seven classes formed the Free Reading group and the Control group. The stratification of students on the following characteristics was done to control some possible confounding factors. As explained above, the homogenous background and school schedules of the subjects mean that extraneous variables are not particularly difficult to control.

In previous studies, researchers have shown their concern to a variety of extraneous variables that might influence the results of extensive reading. The extraneous variables noted in the literature included: teaching methods, teachers, textbooks, teaching and learning facility, class time schedule, experimental mortality, loss of subject, individual students' social background, intelligence and enthusiasm, interest, anxiety and motivation in learning English (Campbell & Stanley, 1966; Tan, 1998).

In the present study, the primary concern was the participants' previous English language proficiency. Ideally, the participants' English language proficiency in

these three groups should be approximately the same at the beginning of the treatments. A great discrepancy of the participants' English language proficiency among the three groups will make it difficult to judge whether the differences in the mean scores are caused by the previous discrepancy or by the treatments. Therefore the participants were first stratified on the basis of their performance on the pre-test. As shown in Table 3.9 below, the stratification result was that the participants in the three groups had approximately equal English language proficiency at the beginning of the program with all of the three groups scoring 53.12 in the pre-test.

Table 3. 9 Descriptive statistics and results of ANOVA from the pre-test

| Sections (Total score) | Group | N | Mean | SD | F | Sig. |
|------------------------|-------|----|-------|------|------|------|
| Grammar (10) | CR | 33 | 3.61 | 1.20 | 0.94 | 0.39 |
| | FR | 33 | 3.97 | 0.95 | | |
| | CG | 33 | 3.91 | 1.28 | | |
| Cloze (30) | CR | 33 | 14.18 | 2.51 | 0.78 | 0.46 |
| | FR | 33 | 13.52 | 2.59 | | |
| | CG | 33 | 13.39 | 3.16 | | |
| Reading 1(22) | CR | 33 | 12.48 | 2.65 | 0.47 | 0.63 |
| | FR | 33 | 12.54 | 3.40 | | |
| | CG | 33 | 11.82 | 3.98 | | |
| Reading 2 (8) | CR | 33 | 5.24 | 1.56 | 0.16 | 0.86 |
| | FR | 33 | 5.18 | 1.57 | | |
| | CG | 33 | 5.39 | 1.64 | | |
| Translation (10) | CR | 33 | 7.03 | 1.59 | 0.66 | 0.52 |
| | FR | 33 | 6.64 | 1.65 | | |
| | CG | 33 | 6.61 | 1.78 | | |
| Writing (20) | CR | 33 | 11.06 | 2.08 | 0.23 | 0.80 |
| | FR | 33 | 11.30 | 2.07 | | |
| | CG | 33 | 11.39 | 2.05 | | |
| Total (100) | CR | 33 | 53.12 | 5.45 | 0.00 | 1.00 |
| | FR | 33 | 53.12 | 4.88 | | |
| | CG | 33 | 53.12 | 6.16 | | |

* The mean difference is significant at the 0.05 level.

*The results of Levene tests show there is no deviation from homogeneity of variance.

Note: Reading 1 was in the form of multiple choice and Reading 2 was short questions and answers.

Given that the pre-test was developed and confined within a school district only, the students' total scores obtained in the Shanghai Senior High School Matriculation Test (SSHMT) were examined. The results are presented in Table 3.10 below.

Table 3. 10 Descriptive statistics of the SSHMT scores of the three groups

| Group | N | Mean | SD | Minimum | Maximum |
|-------|----|--------|------|---------|---------|
| CR | 33 | 570.15 | 7.37 | 554.50 | 586.50 |
| FR | 33 | 570.26 | 7.36 | 554.00 | 579.50 |
| CG | 33 | 570.88 | 8.71 | 554.00 | 588.50 |
| Total | 99 | 570.43 | 7.77 | 554.00 | 588.50 |

As shown in Table 3.11 below, there was no statistically significant difference between each group, which meant that the three groups were comparable in terms of both the English language proficiency and overall learning ability. It also confirmed that the pre-test was valid and reliable, and the three groups established on the basis of the pre-test were comparable.

Table 3. 11 The results of multiple comparisons of the SSHMT scores by Fisher's LSD Test

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| SSHMT Score | CR | FR | -0.11 | 1.93 | 0.96 |
| | | CG | -0.73 | 1.93 | 0.71 |
| | FR | CR | 0.11 | 1.93 | 0.96 |
| | | CG | -0.62 | 1.93 | 0.75 |
| | CG | CR | 0.73 | 1.93 | 0.71 |
| | | R | 0.62 | 1.93 | 0.75 |

*The mean difference is significant at the 0.05 level.

*The results of Levene tests show there is no deviation from homogeneity of variance.

The second concern was gender. It was pointed out by previous researchers that females might be better at language learning than males (Ellis, 1994; Nyikos, 1990). A number of studies show that females have more positive attitudes to learning a second language (L2) than males (Bacon & Finnemann, 1992; Spolsky, 1989). Therefore the three groups were made to comprise exactly the same number of female and male students to minimize the influence of the gender variable, as shown in Table 3.12 below.

Table 3. 12 Genders of the participants in the three groups

| Group | Female Students | Male Students | Total |
|-------|-----------------|---------------|-------|
| CR | 23 | 10 | 33 |
| FR | 23 | 10 | 33 |
| CG | 23 | 10 | 33 |

The third concern was the amount of time the participants spent on English language learning. Time was a crucial factor in the present study. Conditions being similar, the participants who spent more time studying English were more

likely to make bigger progress than those who did less time. In the sample school, the students in different classes adhered to the same curriculum standards and timetable. Every week there were five sessions of forty minutes for in-class teaching. On the surface, there was no doubt that the participants spend approximately the same amount of time studying English. However, it was still possible that different teachers would assign different amount of homework.

The third concern thus gave rise to the fourth concern: the teachers. The teacher variable was one of the most important extraneous variables in the present study. Other conditions being equal, different teachers will produce different teaching results. The first factor that matters is the teaching style and personality of the teacher. Some teachers are able to maintain students' interest until the end of a class, while some others will induce sleep. The second factor that matters was the teachers' teaching experience. Teachers with longer teaching experience tend to gain better understanding of teaching and students, and tend to be better at classroom management. The problem was that it was difficult to quantify how experienced a teacher was. A teacher's working year may be a direct scale to quantify his or her teaching experience. The researcher had three years' teaching experience before the present study began. The average teaching experience of the other six teachers in the same grade, was 7.5 years, 250% longer than that of the researcher. Individually, the most senior teacher had 23 years' teaching experience, and the youngest four years. The calculation and comparison of teaching experience suggests that the researcher in charge of the Complementary Group was weaker than her colleagues in terms of working experience. It is reasonable to say that if extensive reading produces better results than others, it cannot be credited to the researcher's advantage in teaching experience; on the contrary, she was weaker in this aspect.

In addition to the analysis, a further step to minimize the influence of the third and fourth extraneous variables was that the Free Reading group and the Control Group were made to comprise students from every other seven classes instead of a single one, and the composition of the Free Reading group and the Control Group were exactly the same in terms of numbers of students from different classes as illustrated in Table 3.1 in Section 3. 2.

Eventually the final sample consisted of 99 students with approximately equal beginning English language proficiency, 33 in the Complementary Reading group, 33 in the Free Reading group and 33 in the Control Group. Among the 33 students in each group, 23 were girls and 10 were boys. The complementary reading participants were in the same class, and the other 66 participants were from the other seven classes.

3.6 Treatments

Efforts have been made to minimize the effects of the possible extraneous variable. The crucial difference between the three groups was the treatments. The Control group adhered to the standard syllabus and curriculum, and in the two experimental groups, a new extensive reading component was integrated, which is detailed in this section.

As discussed in Chapter 2, there are different ways to define extensive reading. The three major trends in defining extensive reading are: extensive reading as a reading style, extensive reading as an input medium and extensive reading as a teaching approach (Irvine, 2006). The two treatments adopted in the present study reflect the latter two: extensive reading as an input medium and a teaching approach. The first treatment was the free reading treatment, and the latter was the complementary reading treatment.

The reason for setting two experimental treatments was that although extensive reading is now widely accepted as an effective way to improve learners' language proficiency, there is less agreement on the best way of implementing it in the classroom. Administering two different treatments to two experimental groups investigates how extensive reading can be more effectively implemented in this context.

3.6.1 The Free Reading Treatment

Because the free reading participants came from different classes in the same grade, they adhered to the normal curriculum in class and did homework assigned by their respective class English teachers after school. In class, this group of participants was explicitly taught their regular textbook, *Oxford English for Senior*

High School, in class by their class English teachers. Out of class, they read the graded readers when they had time for reading, and were encouraged to read as much as they could. They were advised to keep borrowing books every week, and submit book reports when they returned the books. There was no class time devoted to free reading treatment. The researcher supervised the free reading participants mainly through reading, marking their book reports and checking their borrowing records. If a free reading participant was considered to read less or to be less diligent than their peers, the researcher would have a private talk with the participant, trying to find out why he/she read less or if he/she had met any difficulty that prevented him/her from reading more. The researcher had been trying to encourage them to read more, and helped them in every possible aspect although she was not their class English teacher.

3.6.2 The Complementary Reading Treatment

The teaching of the textbook, *Oxford English for Senior High School*, was a core component in the standard curriculum, the absence of which was strictly prohibited. The same as the Free Reading participants, the Complementary Reading participants were explicitly taught their regular textbook *Oxford English for Senior High School* by their English teacher, (also the present author).

In addition to this similarity, the core difference between the two experimental groups was that extensive reading was a medium of extracurricular input for the Free Reading group while it was integrated into the English curriculum for the Complementary Reading group (CR). The Complementary Reading group shared their reading via in-class discussion and part of the class time was devoted to in-class discussion. To provide a basis for in-class discussion, each title at the same level was marked with a number. For instance, there were 41 titles at Level 1. On Monday, student No.1 read book No.1 and student No. 2 read book No. 2. On Tuesday, student No. 1 read book No. 2, and student No. 2 read book No. 3, and so on. They did not need to approach the book corner; instead, they exchanged their books day by day. For example, student No.1 passed his/her book to student No. 2, and received a book from student No. 33. The name list of the whole class was posted on the wall, so the complementary reading participants knew to whom they pass the books and from whom they received

books. When the Complementary Reading participants moved to Level 3 and Level 4, which required more days to finish, they did not exchange books day by day but exchanged books after they finished reading them.

The Complementary Reading treatment distinguished itself from the procedures of extensive reading implementation proposed by Day and Bamford (1998) in the following important aspects. Firstly, the Complementary Reading treatment allocated no time for in-class reading. The reading was done out of class, before each session began. The complementary reading participants were required to read around 5,000 words per day, which meant they finished reading a Level 1 book in one day, a Level 2 book in one day, a Level 3 book in two days and a Level 4 book in three days. Secondly, the participants were not able to choose what they wanted to read. Thirdly, they were requested to finish reading what was allocated to them and wrote up a brief summary of the content.

There were five sessions of forty minutes for English every week. Before the books arrived, i.e., the official start of the program, five to ten minutes was devoted to issues related to the present study in two to three sessions every week. It included the contrast between the traditional Grammar-Translation Method and extensive reading, introduction of the benefits of extensive reading, introduction of the books and program, training of necessary strategies and skills for extensive reading, demonstration of how to arrange in-class discussion and question and answers to other issues related. This training before the official start of the present study smoothed the transition from the traditional teaching method to a complementary of extensive reading and normal classroom teaching.

After the official start of the program, the frequency of in-class discussion was two or three times a week with ten minutes devoted to discussion each time. If the discussion was heated, the researcher would not interrupt, but let the discussion proceed. In-class discussion was student-oriented and focused on a commonly read story. Every student took turns to arrange and lead the discussion with the student on duty responsible for reporting to the whole class and the researcher about what he or she had read and then organizing the discussion. Since the students were not very good at oral English, students were allowed speak in

Chinese so they could be actively involved without worrying about their oral English proficiency.

At the very beginning, the organizer appeared to be too much of a reporter and volunteer therefore few of the classmates were willing to join the discussion. The first reason was that they had not read the book being reported and discussed. The second reason was that they were not used to this form of classroom teaching as they were used to the teacher talking all through the class. In this situation the researcher played an active role in the discussion, asking or answering some questions to push the discussion. She cooperated with the organizer to show the whole class how the discussion could go on. About one month later, the discussion began to go on well. That was because the complementary reading participants had almost finished reading Level 1 in the past month, which served as a basis for mutual discussion. The second reason was that the researcher had showed them a practical model of discussion. The third reason was probably that they would feel relaxed to express themselves and share what they thought about the books and stories with their peers.

Reading served as a stimulus to discussion. Sometimes the complementary reading participants managed to organize the discussion on their own. In this situation, the discussion was centered on the following topics: (1) the authors, especially the writers well-known to the Chinese, such as Charles Dickens, Sir Conan Doyle and William Shakespeare, etc; (2) some masterpieces which are also well-known to the Chinese, such as *Aladdin and the Enchanted Lamp* and *Robinson Crusoe*, etc; (3) some titles which have been adapted into films and dramas, such as *Titanic*, *The Phantom of the Opera* and *Hamlet*, etc; and (4) some countries and regions introduced in *Factfiles*, such as Scotland, Australia, New Zealand and the USA.

Sometimes, the complementary reading participants raised questions about the background knowledge regarding the author and his/her writing background, major events in a book, foreign governments and educational systems, foreign customs and so on. To answer some of the questions, the researcher talked about the topic explicitly in class. For instance, some students did not understand why

the United Kingdom and France were deadly enemies when they were reading *A Tale of Two Cities* (Level 4). The researcher made careful preparation and delivered a mini lecture on the relationship between the United Kingdom and France in history, which was well-received by students. To some of the questions, the researcher recommended some books for students to read. For example, some students admired Nelson Mandela greatly, but they could not imagine how hard it was to liberate South Africa and abolish apartheid. The researcher recommended that they read *Land of My Childhood* (Level 4) and *Martin Luther King* (Level 3) in the same series to find out how the black people in Africa worked and lived under apartheid, how the black people in America have been fighting against racial discrimination, and how great the two leaders were to make everyone equal. This kind of engagement shown by students is an important aspect of extensive reading and is under-reported in the literature. The discussion broadens students' knowledge of the world and may be seen as the first step in reading across the curriculum. There is also an element of moral training introduced, e.g. the discussion of Mandela. Thus discussions should be part of pedagogy and not just used for research. The empirical research framework permits this blending of teaching and research.

3.6.3 English language teaching in the Control group

The traditional Grammar Translation Method was employed for the control group. This group of participants was explicitly taught their regular textbook, *Oxford English for Senior High School*, in class by their class English teachers. They did not read extensively and instead adhered to the normal syllabus and teaching content, with their performance on the instruments serving the purpose of comparison and contrast with the two experimental groups.

There were no special additional reading materials for the control group. The only reading resource available to them was their regular textbook. In each textbook twelve articles ranging from 300 to 500 words are arranged into six units, with two articles in each unit. The teaching of the twelve articles lasts the academic term, around eighteen teaching weeks.

3.6.4 Monitoring

As has been described above, the complementary reading treatment was an outside-class pleasure reading plus inside-class reading community and the role of the researcher was mainly as group member and counselor. However, as a teacher, it was very important for her to monitor the participants' learning progress and to get feedback from the participants so as to help them to solve some problems and adjust to the program. In the traditional teaching pattern, this role was played mainly through class instruction, question-and-answer type of exercises and tests. Given that there were no follow-up tests in the experimental groups, the following complementary measures were adopted by the researcher to monitor the participants' reading activities and helped to solve the participants' problems.

1. Mark the participants' reports. Having read a book, the participants in both experimental groups were required to write a brief book report about 80-150 words in English. There was no restriction on the content of the report. It might be comments on the theme, the characters, the plots or anything the students felt worth writing.
2. Nearly every day, the researcher had informal talks with the complementary reading participants, trying to identify their reading problems. Due to the full school timetable, the researcher and the complementary reading participants usually talked with each other using the 10-minute break between each session or the 30-minute lunch break. If the researcher identified that some complementary reading participants lagged behind or had some special problems, she would have a private talk with them to offer help and encouragement. At different stages, nearly all complementary reading participants lagged behind and for different reasons. In the first stage the complementary reading participants with relatively lower English language proficiency than their peers showed reluctance and little confidence. They received encouragement and individual instruction on extensive reading from the researcher. In the second phase, some complementary reading participants felt that they did not make as much progress as they had expected from extensive reading. They were losing interest and motivation. The researcher identified their progress both from

their daily study and extensive reading, helping them keep up with other classmates. More than one third of the complementary participants experienced this problem. In order to ensure the program moved along smoothly, the researcher picked up four sentences of each student's book reports, two from the earliest ones, and two from the most recent ones. The two groups of sentences were put together and shown to the whole group in class. By doing this, the complementary reading participants identified their progress both in their reading and writing, and regained confidence and motivation. In the last phase, when the final-term test was approaching, nearly all the complementary reading participants complained about overload of homework. Most of them still could not finish homework until mid-night. In this situation the researcher cut down the daily reading amount, allowing them more time for their examination preparation.

The monitoring of the free reading participants was more difficult. The free reading participants were not in the researcher's class so the researcher's authority and influence on them was understandably less. To talk with the free reading participants in order to encourage them to keep reading and collect their feedback on the program, the researcher needed to talk with their English teacher and class teacher first and then found a suitable time. The talk could not last very long because a long talk would either delay the participants for the subsequent teaching session or their going back home after school. Fortunately, the free reading participants were very considerate. The talk with them was always enjoyable and constructive.

3.6.5 Summary

To summarize, different conditions for English language teaching and learning were created in the three groups as shown in Table 3.13 below:

The different conditions for the three groups are summarized below:

The Complementary Reading group: The teaching of the regular textbook, *Oxford English for Senior High School*, was a core component in the standard curriculum, the absence of which was strictly prohibited. This group of participants was explicitly taught their regular textbook by their English teacher, (also the researcher). However, they did not do any additional

grammar-translation homework after school; instead, extensive reading was their daily homework which was complementary to the classroom teaching, and they read around 5,000 words every day after school.

The Free Reading group: This group of participants was explicitly taught their regular textbook in class by their class English teachers. They did their grammar-translation exercises after school and read out of class only when they had time without specific amount of reading required.

The Control group: This group of participants was explicitly taught their regular textbook in class by their class English teachers. They did grammar-translation exercises after school and did not read extensively. They were taught using typical grammar-translation methodology, and were not exposed to English in their schooling outside of their textbooks.

Table 3. 13 Summary of the treatments in the three groups

| | Control group | FR group | CR group |
|---|---|---|---|
| In-class learning activities | Standard curriculum with the regular textbook <i>Oxford English for Senior High School</i> | | Standard curriculum with the regular textbook <i>Oxford English for Senior High School</i> + 10-minute discussions |
| Out-of-class learning activities/homework | Grammar translation exercises | Grammar translation exercises + free extensive reading | Complementary extensive reading of about 5,000 words per day |

3.7 Challenges in implementation

Extensive Reading has a low profile in mainland China because it is a new approach, especially for Chinese students and teachers. Extensive Reading is significantly different from the traditional approaches they are used to or trained in, and thus there are often challenges in implementing extensive reading in this specific context. The challenges that hindered successful practice of extensive reading observed in this program included: reluctant readers, the change of the role of the teachers, the change of the classroom culture, and students' academic pressure from the Chinese context.

Reluctant readers

In Takase's (2007b) study (see Section 2.5.3), readers who are not interested in reading is the third biggest concern shared by the extensive reading practitioners.

There is a wide range of factors which may account for the students' reluctance to read in English. The first factor may be students' low confidence in English language reading. Reading in English is for them a passive and frightening activity associated with countless unknown words, complex sentence structure and old-fashioned topics before they actually read the first book. The strategies and skills are needed to be cultivated, because they are not automatically concurrent with piles of books or a class library. Moreover, the 1990s rapidly spawned a generation whose free time was filled with video games, comics and cartoons. Any of these colorful visual images are more appealing than black and white books. Even if they are convinced that reading is pleasurable, their main "pleasure" is usually in answering the researcher's or teacher's comprehension questions.

The washback effect of the NMET is another factor contributing to the students' reluctance. In the NMET, the reading comprehension section is composed of three to four short passages with 300 to 500 words (see Section 1.4 for detailed introduction of the NMET; see Cheng, 2008; Cheng & Qi; Li, 1990; Qi, 2005, for its prominence for in the Chinese educational system). Maxim (2002) reported that it was likely that students equated reading in English with reading short passages, which naturally kept them away from book-length reading, nor did they have any intention to approach such a mysterious task. Maxim's report matches our experience that some students were not certain whether their time and effort on reading was cost effective, i.e., whether reading would bring them rewards in standardized English exams. Even after the students have begun their reading, further problems occur. Jackson (2005) noted that her students tended to read much too difficult reading materials, and lack awareness of the reasons for extensive reading, and how to go about it successfully.

When faced with this challenge, the researchers adopted a simple and straightforward solution: starting the program with an easy book. All the participants in this program were advised to begin their Extensive Reading with a starter or a Level 1 book, which they were fully capable of. After finishing reading the first book written in English, they felt proud and confident. The lack of interest in reading, lack of confidence in their reading ability, worry about

much too difficult reading material, in other words, their anxiety about reading in English had been greatly reduced when they finished reading the first book, as they noted in the communication with the researcher.

Furthermore, this program showed the participants that Extensive Reading is very interesting and can initiate various interactions, which interested them greatly. The presence of “discussion” activities also encouraged the reluctant readers. If extensive reading is implemented with a reading community in which participants exchange ideas and thoughts freely, their motivation to read will be strengthened.

The change of teachers' role

The change of the role of the teachers was another challenge facing Extensive Reading practitioners. Traditionally, the role of a teacher in the Chinese context is a ruler of the classroom and speaker of the truth. Their authority is never challenged. When asking questions, teachers usually ask questions for which they are fully prepared (Widdowson, 1990), because it is considered to be the teachers' responsibility to give a standardized answer to the question asked. This belief is further enhanced by the high-stakes exams such as the NMET, in which multiple choices account for more than 60% of the test content. Influenced by the Chinese learning culture, the washback effect of high-stakes exams and teachers' role, the students take it for granted that there is only one answer to any question, and it has seldom dawned on them that there may be many other possible answers. The complementary of extensive reading into curriculum may initiate various interaction and discussion, and accordingly increases teachers' risk of not knowing the standard answers to questions asked by students as the questions may cover history, society, and so on. In this program, to make herself competent in answering students' questions from their reading, the first author had to spend time preparing herself very well. She read all the 235 books used in this program, and read extensively about the background information of the 235 titles.

In addition to the change of teachers' role, extensive reading brought extra burdens to teachers. While many of the courses require little preparation outside of class, extensive reading is one of the most demanding tasks. It takes the

English teachers up to five hours a week of outside preparation (Robb, 2001). In the present study, it took the researcher much more than five hours per week for the preparatory work. In addition to all the book management work, the researcher prepared for in-class talk, motivated the reluctant students and tried to maintain their interest in every possible way.

The researcher felt great pressure, and was busier than ever before compared with her past working experience. However, she felt that it was a worthwhile enterprise to establish the benefits of extensive reading in Shanghai secondary high schools, exposing her students to a wonderful world. She was more experienced in extensive reading after the program, and was willing to pursue it further in the future. As Hill (1997b) concludes, “The books are there waiting to be used. The students are there waiting to read them. The work of bringing them together is very worthwhile” (p. 20).

The change of classroom culture

The work of bringing the books and students together is very worthwhile indeed, but not the end of the challenges. The change of both teachers’ and students’ roles led to the change of classroom culture or the way the teachers and students interacted. In the past if the students wanted to ask questions in class, they needed to raised hands and waited for the teacher’s permission before making a sound; if they were asked questions, they needed to stand up to show respect to the teachers. In this manner, the interaction between students and teachers was highly controlled, let alone the interaction between student and student. In this program, the students were encouraged to voice their opinions freely as they wished. Coupled with the new freedom was the chaos of the classroom. It was not easy for teenagers to achieve the balance between heated discussion and observations of class rules. It took some time for both the researcher and students to get used to the new change and worked out a solution to the chaos: you can speak, but be quiet while someone else was speaking.

Students’ academic pressure from the Chinese context

The last, but biggest factor hindering extensive reading was not the reluctant readers, the change of the teachers’ role or the new interaction but the heavy

learning burden. Since the 2008/2009 academic year, the Shanghai Board of Education has increased the number of core subjects for senior high schools from five to nine. Before 2008, the Shanghai senior high school students were tested on Chinese, math, English, physics and chemistry, while after 2008, the Senior One students were required to take geography and computer tests, and the Senior Two students had to take history and biology tests. The change not only meant two more subjects and two more tests, but it meant much more homework and greater demand on students' academic performance. The Board claimed that they were trying to enrich students' knowledge in science and art, but the direct consequence of the change, in the researchers' eye, was that the students were more tired than ever before. Most of the students complained that they were unable to go to bed until mid-night, and had to get up at six a.m. or even earlier. In such an unpromising situation, how could we expect the students to read in a good mood, or expect them to read as more as possible when their sleeping time was only 5 to 6 hours per day? Under such a circumstance, introducing extensive reading into such classes of students with still more words to read and book reports to write was understandably unwelcome. Although claimed to be reading for pleasure, extensive reading in this situation was not likely to be a pleasurable activity.

3.8 Time frame

The whole program was implemented in the following main steps: (a) design the program, (b) set up the experimental groups, (c) prepare teaching, and (d) employ assessment tools. The program outlined below applied exclusively to the experimental groups as presented in Table 3.14 below.

Table 3. 14 Timeframe of this school-based extensive reading program in 2010/11

| Group | 2010.8 | 2010.9-2011.1 | 2011.1 | 2011.2-2011.6 | 2011.6 | 2012.6 |
|-------|----------|---|----------|---|-----------|-------------------|
| CR | Pre-test | Normal classroom instruction. Complementary reading treatment. | Mid-test | Normal classroom instruction. Complementary reading treatment. | Post-test | Delayed post-test |
| FR | Pre-test | Normal classroom instruction. Free reading treatment. | Mid-test | Normal classroom instruction. Free reading treatment. | Post-test | Delayed post-test |
| CG | Pre-test | Normal classroom instruction. No extra reading. | Mid-test | Normal classroom instruction. No extra reading | Post-test | Delayed post-test |

4 Findings and discussion

The primary purpose of this research undertaking was to investigate the possible significance of extensive reading in Shanghai senior high schools. The free reading treatment investigated the benefits of reading as voluntary additional input to the normal curriculum, while the complementary reading treatment attempted to compare the effect of extensive reading with that of the grammar and translation exercises in the standard curriculum. It hypothesized that extensive reading could positively influence the students in the following proficiency areas: reading, listening and writing, and the two language knowledge components, namely, vocabulary and grammar. In addition, this study investigated how extensive reading can be effectively implemented in this research context, and was examined through the use of two different extensive reading methodologies for the two different experimental groups.

4.1 Data analysis

The tests and reading record data were first coded to be computer-readable and then entered into the computer via a data-entry program to construct three data files: one for the pre-, mid-, post- and delayed post-tests, one for the pre- and post-program EPER Placement Test, and another for the reading record. Statistical analyses of the data were performed by Statistical Package for the Social Science (SPSS) Version 17.0.

The pre-, mid-, post- and delayed post-tests data were entered into an excel worksheet and categorized into listening, grammar, vocabulary, cloze, reading, translation and writing in light of the availability and sequence of these components in the tests. Analysis of Variance (ANOVA) was employed to examine whether there is a main effect of a particular variable on the three groups. If so, post hoc Fisher's LSD test (see Section 4.3 for detailed introduction of the Fisher's LSD test) was conducted to carry out multiple comparisons between groups to examine whether there was statistically significant difference between groups on specific language knowledge components and language skills. The pre- and post-EPER Placement Test results were also entered into an excel worksheet and categorized into pre-test original

score, pre-test standard score, pre-test EPER level, post-test original score, post-test standard score, post-test EPER level and progress in terms of EPER levels between two tests. A paired-sample *t*-test was conducted to examine the data, given that two sets of data from the same EPER Placement Test (i.e., pre- and post-tests) were compared.

The complementary and free reading participants' individual reading records were first entered into 66 separate excel worksheets according to their borrowing records and book reports. The researcher had been keeping an accurate record of the participants' reading, including the number, the date on which the participants read the books, the titles, the levels of the books, the word count of each book, the exact time students spent on each book, the average reading speed calculated by word count/time, and students' rating of the book and any comments or suggestions on the book. Table 4.1 below presents the format of the participants' reading record.

Table 4. 1 Format of participants' reading record

| No. | Year | Date | Title | Level | Genre | Word Count | Time (m) | Speed (w/m) | Rate | Comments & Suggestions |
|-----|------|-------|-------------------|-------|----------|------------|----------|-------------|------|------------------------|
| 15 | 2010 | 11.28 | A Christmas Carol | 3 | Classics | 10385 | 90 | 115.4 | 4 | |

Note: 1. m=minutes, w/m=word per minute

2. The students rated the titles on a five-point Likert scale, with 1 the lowest, meaning very dissatisfactory, and 5 the highest, meaning highly satisfactory.

After the present study ended, the researcher printed out the individual reading records and distributed the records to the participants. The participants were asked to first check the printed records with their own reading reports, and second to make revisions on their ratings of titles and add comments and suggestions for individual titles if they wished. With the recheck of individual recording records finished and all the errors corrected, the researcher entered the individual reading records into two excel worksheets: reading record of the Complementary Reading group and reading record of the Free Reading group. The two summarized records provided the following information: numbers of books read by individual participants, numbers of books read by the group, numbers of books read by individual participants at different levels, numbers of books read by the group at different levels, reading amount of individual

participant, reading amount of the two groups, frequency of the genres read by the participants, preference for different genres by the participants, reading speeds of individual students when reading at different levels, average reading speeds of the whole group when reading at different levels, average ratings of a title given by the participants who had read the title.

4.2 Basic information of students' extensive reading

This program based on extensive reading commenced in September, 2010, and ended in June, 2011. As illustrated in Table 4.2, the participants in the Complementary Reading group read 3,660 books in total, 110.9 books per participant on average, approximately equal to 1,144,279 words per student in the whole program. Compared with their complementary reading peers, the participants in the Free Reading group read much less. According to their individual reading records, they read 853 books in total, 25.9 books per student on average. The average reading amount of the Free Reading group was 226,732 words, one fifth of that of the Complementary Reading group.

Table 4. 2 Detailed information of participants' extensive reading

| Group | Semester | Phase | Le0 | Le1 | Le2 | Le3 | Le4 | Le5 | Le6 | Total | Reading amount |
|-------|----------|-----------|------|------|------|------|------|-----|-------|-----------|----------------|
| CR | 1 | 1 | / | 9.4 | 4.2 | 2.5 | 0.2 | / | / | 16.3 | 115,078 |
| | | 2 | / | 16.7 | 13.4 | 8.8 | 0.9 | 0.4 | 0.1 | 40.3 | 303,786 |
| | | Sub-total | / | 26.1 | 17.6 | 11.3 | 1.1 | 0.4 | 0.1 | 56.6 | 418,864 |
| | 2 | 1 | / | 3.6 | 12.1 | 7.6 | 8.4 | 3.8 | 0.1 | 35.6 | 397,917 |
| | | 2 | / | / | 0.1 | 2.4 | 14.3 | 1.8 | 0.1 | 18.7 | 327,498 |
| | | Sub-total | / | 3.6 | 12.2 | 10.0 | 22.7 | 5.6 | 0.2 | 54.3 | 725,415 |
| Total | | / | 29.7 | 29.8 | 21.3 | 23.8 | 6.0 | 0.3 | 110.9 | 1,144,279 | |
| FR | 1 | 1 | 1.3 | 3.9 | 2.2 | 0.9 | 0.2 | / | / | 8.5 | 46,651 |
| | | 2 | / | 1.2 | 2.3 | 1.5 | 0.7 | 0.4 | 0.2 | 6.3 | 65,212 |
| | | Sub-total | 1.3 | 5.2 | 4.5 | 2.4 | 0.9 | 0.4 | 0.2 | 14.8 | 111,863 |
| | 2 | 1 | / | 1.8 | 2.1 | 2.0 | 1.2 | 0.4 | 0.3 | 7.9 | 75,977 |
| | | 2 | / | 0.3 | 1.2 | 1.2 | 0.3 | 0.1 | 0.1 | 3.2 | 38,892 |
| | | Sub-total | / | 2.1 | 3.3 | 3.2 | 1.5 | 0.5 | 0.4 | 11.1 | 114,869 |
| Total | | 1.3 | 7.3 | 7.8 | 5.6 | 2.4 | 0.9 | 0.6 | 25.9 | 226,732 | |

Note 1: Le0=Level 0, Le1=Level 1, etc.

Note 2: The word count of every playscript is not available. In this study, the word count of every playscript is set as 5,000.

The circulation rule for the Complementary Reading group was a mixture of compulsory reading and self-selected reading. Given that the complementary reading participants discussed their reading in class and needed to exchange their books with classmates every day, all the complementary reading participants must read at the same level from Monday to Thursday. In the first half of the first

semester they mainly read Level 1 books, and in the second half Level 2 books. In the first half of the second semester they read Level 3 books and the second half Level 4 books. This helps to explain why the final distribution of the levels of the Complementary Reading group was approximately equal. From Friday to Sunday, and during long holidays, such as the National Day, the Dragon Boat Festival, the May Day holiday and the winter holiday, when the complementary reading participants would have more time to read, they were allowed to choose the books that interested them. The chance of free choice was particularly welcome by the participants who were willing to read books at higher levels and the participants who had special interest in a particular genre.

For the purpose of controlling the whole situation, a supplementary rule was issued to ensure that all the complementary reading participants read approximately the same amount. Based on the rule that every complementary reading participant read about 5,000 words per day, their reading amount during Friday, Saturday and Sunday should be around 15,000 words. 15,000 words can be converted to 3 Level 1 books, 2 Level 2 books, 1.5 Level 3 books, and 1 Level 4 book. The complementary reading participants may borrow 3 Level 1 books, 2 Level 2 books with 1 Level 1 book, a Level 1 book and a Level 3 book, or a Level 4 book for the three days. If wanting to try Level 5 and Level 6 books, they were allowed to keep a Level 5 and Level 6 book for two to four weeks depending on the word count and his/her own reading progress. This rule was agreed by the complementary reading participants because it comprised fairness and their right to choose. In the four-week winter holiday, the 235 books were normally distributed to within the Complementary Reading group. Every complementary reading participant may check out one book from Starter to Level 4 each, and they may choose either a Level 5 or a Level 6 as there were not enough books for everyone at Level 5 and Level 6. This distribution did not mean that they should finish reading what they had; instead, they were encouraged to visit each other at home with the books, and exchanged books. During winter holiday, they were required to finish reading at least 70,000 words.

In comparison, the free reading participants were granted more freedom, as they were free to choose what they wanted to read throughout the program and no

reading amount was required. Interestingly their choices of levels were approximately the same as their complementary reading peers. It confirmed the results of the pre-test from another perspective that the reading proficiency of the students in the two experimental groups was roughly equal.

The difference between the reading amount of the first and second term is also noticeable. Both the Complementary Reading group and the Free Reading group read more in the second term than they did in the first term, especially the Complementary Reading group, who read around 300,000 words more in the second term. The explanation was that the second term was longer than the first term in the present study. In mainland China a whole academic year was divided into two academic terms. The beginning of the first academic term is fixed on September 1st, except when September 1st falls on a Saturday or Sunday. In this situation the new semester in turn is scheduled to begin on September 2nd or 3rd. The termination of the second academic term is scheduled on June 30th, by which time all the tests will have been administered and the test results finalized, because the summer holiday spans two months, from the beginning of July to the end of August. In comparison, the termination of the first academic term and the commencement of the second academic term are not fixed but decided by the Chinese Lunar New Year. Usually the first academic term ends around one week ahead of the Chinese Lunar New Year, and the second academic term commences about two weeks after the New Year. In the 2010/11 academic year, the Chinese Lunar New Year was on January 23rd, and thus the first academic year ended on January 16th, 2011. The whole first term lasted four months and a half, from September 1st 2010 to January 16th, 2011.

Because of the preparatory work of the present study at the beginning of the first term, including numbering the books, giving orientations on extensive reading and training the participants on extensive reading, which took two weeks, and the students' military training in the army, which took one week, the actual time the participants spent on extensive reading was approximately three months and a half, i.e. 14 weeks. The reading amount of the first term covered what the participants read from the beginning of the program to the end of the first term, by which time the mid-test was administered. The second academic term

commenced on February 14th and ended on June 30th, which also lasted four months and a half. Given that all the complementary reading participants and most of the free reading participants kept reading during the four-week winter holidays, the reading done during winter holidays was counted into the reading amount of the second term. Thus all the participants did extensive reading for approximately twenty weeks in the second term, six weeks more than they did in the first term. At the same time, it was probable that the participants read faster in the second term than in the first because they had become more skilled after their extensive reading experience in the first term.

In addition to reading amount and distribution of levels, the students' preference for different genres was calculated. The frequency statistics in Table 4.3 below were listed according to the taste of the Free Reading group since the students in this group had unlimited access to the books and they also had the right to choose. It was considered that the frequency statistic of the Free Reading group would be more objective to reflect the students' preferences for genres than that of the Complementary Reading group, because the complementary reading participants were less lucky in this aspect as they were deprived of the right to choose what they wanted to read due to their treatment.

As shown in Table 4.3 below, *Human Interest*, *Thriller or Adventure*, *Fantasy or Horror* were the favorite among the free reading participants because these three genres were borrowed more frequently than the other six. *Playscripts* and *World Stories* seemed to be much less attractive. Very few participants showed their interest in them as evidenced by the borrowing record of the Free Reading group. It could be assumed that more than half of the free reading participants did not touch or read *World Stories* because the total frequency of borrowing *World Stories* was only 12, 0.36 for each participant on average. The result was not a surprise. It was understandable that the teenagers were more likely to be attracted by stimulating stories and exciting plots. This explains why they were found not to choose non-fiction titles, especially stories set in different social and cultural backgrounds.

Table 4. 3 Frequency of genres read by the Free Reading group and Complementary Reading group

| Genre | The Free Reading group | | The Complementary Reading group | |
|--------------------|------------------------|-------|---------------------------------|-------|
| | Frequency | Order | Frequency | Order |
| Human Interest | 154 | 1 | 592 | 1 |
| Thriller/Adventure | 141 | 2 | 451 | 5 |
| Fantasy/Horror | 133 | 3 | 479 | 4 |
| Crime/Mystery | 107 | 4 | 329 | 7 |
| Classics | 106 | 5 | 509 | 2 |
| Factfiles | 96 | 6 | 506 | 3 |
| True Stories | 82 | 7 | 385 | 6 |
| Playscripts | 22 | 8 | 208 | 8 |
| World Stories | 12 | 9 | 201 | 9 |
| Total | 853 | / | 3660 | / |

To return from the discussion about genres, the recheck of the participants' individual reading record showed that the number of books read and the word count of individual participants varied. In the Complementary Reading group, the participant who read most finished reading 141 books, 1,297,610 words, while the one who read least read 98 books, 959, 541 words as shown in Table 4.4 below.

Table 4. 4 Quantification of participants' extensive reading

| Group | Item | N | Minimum | Maximum | Mean | SD |
|-------|-----------------|----|-----------|------------|--------------|----------|
| CR | Number of Books | 33 | 98.00 | 141.00 | 113.70 | 12.23 |
| | Word Count | 33 | 959541.00 | 1297610.00 | 1,144,278.52 | 70180.95 |
| FR | Number of Books | 33 | 23.00 | 28.00 | 25.85 | 1.46 |
| | Word Count | 33 | 161980.00 | 416634.00 | 226731.97 | 60409.28 |

A difference was found in the Complementary Reading group, who read approximately equal amount per day as required. It was necessary to identify how the difference emerged. The explanation was that some capable readers in the Complementary Reading group requested to read books at higher levels than the assigned ones. For instance, if the whole group were reading at Level 2, they asked to read at Level 3 as some of them would be willing to read at their $i+1$ level or some of them considered that Level 2 was much too easy for them. Thus the absolute number of books was indicative of how many books a student had actually read, but it did not mean that the actual number of running words of 98 books was less than that of 141 books. As discussed in Chapter 2, the counting unit *book* is a vague concept. No strict guideline has been issued to specify what constitutes "a book". Taking the Oxford Bookworms (OBW) series read by students in this project as an example, the word count of an OBW Level 1 varies

from 5,000 to 7,000, while that of an OBW Level 6 is between 25,000 and 35,000. Thus the actual reading amount may vary enormously among participants, and in the long term the gap could be accumulated to a huge figure. Compared with that in the Complementary Reading group, the situation in the Free Reading group was less diversified within the group. There was no obvious difference among the group members in terms of the number of books read, while the actual reading amount ranged from 161,980 to 416,634, a 250,000 word gap.

Although the above explanation accounted for the different book numbers, there was indeed a gap between the actual reading amounts. To examine whether there was a difference between the actual reading amounts within the same group, the individual reading amount was subject to the One-Sample *T*-test. As shown in Table 4.5, the One-Sample *T*-test showed that although there were differences between participants' actual reading amounts, the difference was not statistically significant ($t=.000$, $df=32$, $sig=1.00$).

Table 4. 5 Reexamination of participants' reading amount

| Group | Test Value | One-Sample Test | | |
|-------|------------|-----------------|----|-----------------|
| | | t | df | Sig. (2-tailed) |
| CR | 1141248 | .000 | 32 | 1.00 |
| FR | 226732 | .000 | 32 | 1.00 |

*The mean difference is significant at the 0.05 level.

Before beginning to analyze the tests outcomes, we have confirmed that there was no statistically significant difference of the actual reading amount among the participants within the same group, which meant that all the participants in the Complementary Reading group and the Free Reading group had read approximately the same within their own groups. Thus the tests outcomes were attributed to the two extensive reading treatments. It was not because of the different reading amounts among the students that the tests produced different results.

4.3 The test results

4.3.1 The pre-, mid-, post- and delayed post-tests

Prior to presenting the results of the pre-, mid- and post-tests, it should be noted

here that the results of the subsections in the pre-, mid- and post-tests are presented in light of their availability and sequences in the test papers (see Section 3.3 for the detailed introduction of the pre-, mid- and post-tests). Another point worth noting here is that the mid-, post- and delayed post-tests are in strict accordance with the NMET in terms of test format (see Section 1.4 for the detailed introduction of the NMET). The sequence of the seven subsections in the mid- and post- and delayed post-tests is listening, grammar, vocabulary, cloze, reading, translation and writing (see Appendix B for the mid-test and Appendix C for the post-test). The difference between the pre-test and these three tests is that there is not a listening or vocabulary section in the pre-test. The sequence of the five subsections in the pre-test is grammar, cloze, reading, translation and writing (see Appendix A).

In August, 2010, all the participants of the present study took the pre-test, which was the stratification test. The test results were collected and analyzed using SPSS 17.0. Given that we are going to deal with a nominal variable (“Group”) and a scale variable (“Score”), the One-Way Analysis of Variance (ANOVA) test is advisable.

There are three assumptions upon which the one-way ANOVA is based:

1. The scores of three or more independent groups are to be compared on a scale variable.
2. The scale variable is approximately normally distributed.
3. The spread (variances) of the scores for the groups on the scale variable are roughly equal.

(Connolly, 2007, p. 208-209)

The first assumption has been met as we are comparing the scores of the two experimental groups and one control group, and it has been checked that the scale variable (scores) is approximately normally distributed. The third assumption, i.e., the homogeneity of the variance of the dependent variable (score) between groups, is particularly important to ANOVA. Several tests are available for testing this assumption, for example, Levene, Cochran’s C and Bartlett-Box (Hair, et al., 1998). The Levene Test is employed for the present

study to examine the homogeneity of the variance as it is an integral part of the one-way ANOVA procedure in SPSS. As such we can go straight on to run the statistical test. The descriptive statistics and results of ANOVA from the pre-test are presented in Table 4.6 below.

Table 4. 6 Descriptive statistics and results of ANOVA from the pre-test

| Sections (Total score) | Group | N | Mean | SD | F | Sig. |
|------------------------|-------|----|-------|------|------|------|
| Grammar (10) | CR | 33 | 3.61 | 1.20 | 0.94 | 0.39 |
| | FR | 33 | 3.97 | 0.95 | | |
| | CG | 33 | 3.91 | 1.28 | | |
| Cloze (30) | CR | 33 | 14.18 | 2.51 | 0.78 | 0.46 |
| | FR | 33 | 13.52 | 2.59 | | |
| | CG | 33 | 13.39 | 3.16 | | |
| Reading1(22) | CR | 33 | 12.48 | 2.65 | 0.47 | 0.63 |
| | FR | 33 | 12.54 | 3.40 | | |
| | CG | 33 | 11.82 | 3.98 | | |
| Reading2 (8) | CR | 33 | 5.24 | 1.56 | 0.16 | 0.86 |
| | FR | 33 | 5.18 | 1.57 | | |
| | CG | 33 | 5.39 | 1.64 | | |
| Translation (10) | CR | 33 | 7.03 | 1.59 | 0.66 | 0.52 |
| | FR | 33 | 6.64 | 1.65 | | |
| | CG | 33 | 6.61 | 1.78 | | |
| Writing (20) | CR | 33 | 11.06 | 2.08 | 0.23 | 0.80 |
| | FR | 33 | 11.30 | 2.07 | | |
| | CG | 33 | 11.39 | 2.05 | | |
| Total (100) | CR | 33 | 53.12 | 5.45 | 0.00 | 1.00 |
| | FR | 33 | 53.12 | 4.88 | | |
| | CG | 33 | 53.12 | 6.16 | | |

* The mean difference is significant at the 0.05 level.

Note: Reading 1 was in the form of multiple choice and Reading 2 was short questions and answers.

Before considering the results of ANOVA presented in Table 4.6 above, we need to check that the variances of the scores for the three groups are roughly the same by the Levene's test as introduced above. If the significance of the Levene's test is greater than 0.05, it can be assumed that the variances of the scores for the three groups are roughly the same; otherwise, if the significance of the Levene's test is 0.05 or less, we have to conclude that the variance of the scores for the three groups are not roughly equal. In this situation, an alternative statistical test (for the present study the Krusal-Wallis test) will be employed for data analysis.

As shown in Table 4.7 below, results of the Levene's test of each subsection as well as the total score is greater than 0.05. It can be concluded that the variances of the scores for the three groups are roughly the same, which means the three

assumptions underpinning AVNOA have been met and ANOVA is suitable for analysis the pre-test results.

Table 4. 7 Test of homogeneity of variances from the pre-test

| Sections (Total score) | Levene Statistic | df1 | df2 | Sig. |
|------------------------|------------------|-----|-----|------|
| Grammar (10) | 1.128 | 2 | 96 | .328 |
| Cloze (30) | .179 | 2 | 96 | .836 |
| Reading 1(22) | 1.705 | 2 | 96 | .187 |
| Reading 2 (8) | .141 | 2 | 96 | .868 |
| Translation (10) | 1.288 | 2 | 96 | .281 |
| Writing (20) | .051 | 2 | 96 | .950 |
| Total (100) | .660 | 2 | 96 | .519 |

* The mean difference is significant at the 0.05 level.

As shown in Table 4.6, the Complementary Reading group was better at cloze and translation, while the Free Reading group did well in grammar and reading. Questions and Answers (Q & A) and writing were the Control group's areas of expertise. The three groups had their own strong and weak points in different sections or proficiency areas. However, results of ANOVA show that there is no main effect for the groups as a factor in the participants' performance on the pre-test. As discussed in Chapter 3, the pre-test served to establish three comparable groups with approximately equal English language proficiency at the beginning of the present study. The descriptive statistics and results of ANOVA from the pre-test presented in Table 4.6 confirm that there is no main effect for the groups as a factor in the participants' performance on the pre-test, which shows that the three groups entered the present study with approximately equal English language proficiency as well as approximately equal proficiency in grammar, cloze, reading, translation and writing. The results of the multiple comparisons by Fisher's LSD test (detailed below) confirm that there is no statistically significant difference between the groups in the sections covered in the pre-test.

The mid-test was administered in January, 2011, which was also the first final-term English test (see Section 3.3.2 for the making procedures of the mid-test, see also Appendix B for the test paper of the mid-test). As shown in Table 4.9 below, there is a significant main effect for the treatments as a factor in the vocabulary and reading sections. There is also a significant main effect for the treatments as a factor in the total scores of the three groups. Judging by mean

scores, the two reading groups outperformed the Control group in listening, grammar, cloze and writing, but the ANOVA test did not detect a main effect for the treatments as a factor in the difference in these sections.

Table 4. 8 The results of multiple comparisons by the Fisher's LSD test in the pre-test

| Section (Total score) | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|-----------------------|-----------|-----------|-----------------------|------------|------|
| Grammar (10) | CR | FR | -0.36 | 0.28 | 0.20 |
| | | CG | -0.30 | 0.28 | 0.29 |
| | FR | CR | 0.36 | 0.28 | 0.20 |
| | | CG | 0.06 | 0.28 | 0.83 |
| | CG | CR | 0.30 | 0.28 | 0.29 |
| | | FR | -0.06 | 0.28 | 0.83 |
| Cloze (30) | CR | FR | 0.67 | 0.68 | 0.33 |
| | | CG | 0.79 | 0.68 | 0.25 |
| | FR | CR | -0.67 | 0.68 | 0.33 |
| | | CG | 0.12 | 0.68 | 0.86 |
| | CG | CR | -0.79 | 0.68 | 0.25 |
| | | FR | -0.12 | 0.68 | 0.86 |
| Reading 1(22) | CR | FR | -0.06 | 0.83 | 0.94 |
| | | CG | 0.67 | 0.83 | 0.43 |
| | FR | CR | 0.06 | 0.83 | 0.94 |
| | | CG | 0.73 | 0.83 | 0.39 |
| | CG | CR | -0.67 | 0.83 | 0.43 |
| | | FR | -0.73 | 0.83 | 0.39 |
| Reading 2 (8) | CR | FR | 0.06 | 0.39 | 0.88 |
| | | CG | -0.15 | 0.39 | 0.70 |
| | FR | CR | -0.06 | 0.39 | 0.88 |
| | | CG | -0.21 | 0.39 | 0.59 |
| | CG | CR | 0.15 | 0.39 | 0.70 |
| | | FR | 0.21 | 0.39 | 0.59 |
| Translation (10) | CR | FR | 0.39 | 0.41 | 0.34 |
| | | CG | 0.42 | 0.41 | 0.31 |
| | FR | CR | -0.39 | 0.41 | 0.34 |
| | | CG | 0.03 | 0.41 | 0.94 |
| | CG | CR | -0.42 | 0.41 | 0.31 |
| | | FR | -0.03 | 0.41 | 0.94 |
| Writing(20) | CR | FR | -0.24 | 0.51 | 0.63 |
| | | CG | -0.33 | 0.51 | 0.51 |
| | FR | CR | 0.24 | 0.51 | 0.63 |
| | | CG | -0.09 | 0.51 | 0.86 |
| | CG | CR | 0.33 | 0.51 | 0.51 |
| | | FR | 0.09 | 0.51 | 0.86 |
| Total(100) | CR | FR | 0.00 | 1.53 | 1.00 |
| | | CG | 0.00 | 1.53 | 1.00 |
| | FR | CR | 0.00 | 1.53 | 1.00 |
| | | CG | 0.00 | 1.53 | 1.00 |
| | CG | CR | 0.00 | 1.53 | 1.00 |
| | | FR | 0.00 | 1.53 | 1.00 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Table 4. 9 Descriptive statistics and results of ANOVA from the mid-test

| Sections (Total score) | Group | N | Mean | SD | F | Sig. |
|---------------------------|-------|----|-------|------|------|------|
| Listening (20) | CR | 33 | 15.24 | 2.24 | 0.28 | 0.76 |
| | FR | 33 | 15.42 | 2.00 | | |
| | CG | 33 | 15.00 | 2.68 | | |
| Grammar (20) | CR | 33 | 14.15 | 2.62 | 2.16 | 0.12 |
| | FR | 33 | 15.15 | 2.55 | | |
| | CG | 33 | 13.97 | 2.27 | | |
| Vocabulary (9) | CR | 33 | 8.33 | 1.19 | 5.46 | 0.01 |
| | FR | 33 | 8.18 | 1.18 | | |
| | CG | 33 | 7.36 | 1.45 | | |
| Cloze (10) | CR | 33 | 7.79 | 1.27 | 0.58 | 0.56 |
| | FR | 33 | 7.58 | 1.60 | | |
| | CG | 33 | 7.36 | 1.87 | | |
| Reading (16) | CR | 33 | 12.21 | 2.51 | 3.69 | 0.03 |
| | FR | 33 | 12.24 | 2.06 | | |
| | CG | 33 | 10.88 | 2.39 | | |
| Translation (15) | CR | 33 | 11.92 | 1.17 | 1.53 | 0.22 |
| | FR | 33 | 11.68 | 1.62 | | |
| | CG | 33 | 11.15 | 2.47 | | |
| Writing (10) | CR | 33 | 5.67 | 0.70 | 2.43 | 0.09 |
| | FR | 33 | 6.11 | 1.00 | | |
| | CG | 33 | 5.59 | 1.29 | | |
| Total (100) | CR | 33 | 75.32 | 7.75 | 4.04 | 0.02 |
| | FR | 33 | 76.36 | 6.16 | | |
| | CG | 33 | 71.32 | 8.72 | | |

* The mean difference is significant at the 0.05 level.

Table 4.10 below confirms that the variance of the mid-test scores for the three groups are roughly equal before we consider the results of ANOVA from the mid-test.

Table 4. 10 Test of homogeneity of variances from the mid-test

| Section (Total score) | Levene Statistic | df1 | df2 | Sig. |
|-----------------------|------------------|-----|-----|------|
| Listening (20) | 1.583 | 2 | 96 | .211 |
| Grammar (20) | .352 | 2 | 96 | .704 |
| Vocabulary (9) | 2.211 | 2 | 96 | .115 |
| Cloze (10) | 3.008 | 2 | 96 | .054 |
| Reading (16) | .537 | 2 | 96 | .586 |
| Translation (15) | 1.667 | 2 | 96 | .194 |
| Writing (10) | .858 | 2 | 96 | .427 |
| Total (100) | 2.026 | 2 | 96 | .137 |

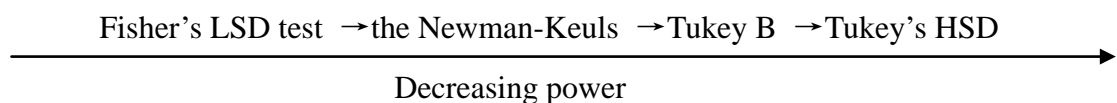
* The mean difference is significant at the 0.05 level.

Results of ANOVA presented in Table 4.9 above have shown that there was a significant main effect for the treatments as a factor in vocabulary, reading and total score. In this situation, a post hoc test would be conducted to identify the location of the statistically significant differences. In the present study, sizes of the three groups were exactly equal, i.e., thirty-three participant in each group. A

for equal group sizes were considered: Tukey's Honestly Significant Difference (HSD) test, Tukey B test on ordered means, Fisher's Least Significant Difference (Fisher's LSD) test, the Newman-Keuls (Neuman-Keuls) test on ordered means and Dunnett test. The Fisher's LSD test was employed and the other three tests were excluded for the following reasons:

1. The Fisher's LSD test is particularly optimal for a study where 3 means are to be compared, because its critical t for significance is not affected by the number of groups (Xu, 2005). In this study, three means, namely the means of the Complementary Reading group, the Free Reading group and the Control group, are to be compared, which makes the Fisher's LSD test a more preferred one among the possible choices.
2. The Fisher's LSD test has the strongest statistical power among the four tests, as illustrated in Figure 4.1 below. The power of a statistical test is defined as its ability to detect a difference or relationship if one does actually exist in the population as a whole (Leonard, 2010). The purpose of the data analysis is to detect the difference among the three groups exactly. A test that has the strongest statistical power, namely the Fisher's LSD test, is optimal for this study.

Figure 4.1 Statistical powers of the post hoc tests



As shown in Table 4.11 below, there is a statistical significance between the two experimental groups and the control group in the vocabulary section. The difference between the two experimental groups in this section is not statistically significant. This situation also applies to reading and total score. In these two sections, the two experimental groups without exception outperformed the control group significantly, and the difference between the two experimental groups in these two sections is not statistically significant. Another point worth

noting is that the Fisher's LSD test detects a statistically significant difference between the Free Reading group and the Control group in the writing section ($p=0.04$), although the results of ANOVA presented in Table 4.9 did not show a significant main effect for the treatments as a factor in writing.

Table 4. 11 The results of multiple comparisons by the Fisher's LSD test in the mid-test

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| Listening | CR | FR | -0.18 | 0.57 | 0.75 |
| | | CG | 0.24 | 0.57 | 0.67 |
| | FR | CR | 0.18 | 0.57 | 0.75 |
| | | CG | 0.42 | 0.57 | 0.46 |
| | CG | CR | -0.24 | 0.57 | 0.67 |
| | | FR | -0.42 | 0.57 | 0.46 |
| Grammar | CR | FR | -1.00 | 0.61 | 0.11 |
| | | CG | 0.18 | 0.61 | 0.77 |
| | FR | CR | 1.00 | 0.61 | 0.11 |
| | | CG | 1.18 | 0.61 | 0.06 |
| | CG | CR | -0.18 | 0.61 | 0.77 |
| | | FR | -1.18 | 0.61 | 0.06 |
| Vocabulary | CR | FR | 0.15 | 0.32 | 0.63 |
| | | CG | 0.97 | 0.32 | 0.00 |
| | FR | CR | -0.15 | 0.32 | 0.63 |
| | | CG | 0.82 | 0.32 | 0.01 |
| | CG | CR | -0.97 | 0.32 | 0.00 |
| | | FR | -0.82 | 0.32 | 0.01 |
| Cloze | CR | FR | 0.21 | 0.39 | 0.59 |
| | | CG | 0.42 | 0.39 | 0.28 |
| | FR | CR | -0.21 | 0.39 | 0.59 |
| | | CG | 0.21 | 0.39 | 0.59 |
| | CG | CR | -0.42 | 0.39 | 0.28 |
| | | FR | -0.21 | 0.39 | 0.59 |
| Reading | CR | FR | -0.03 | 0.57 | 0.96 |
| | | CG | 1.33 | 0.57 | 0.02 |
| | FR | CR | 0.03 | 0.57 | 0.96 |
| | | CG | 1.36 | 0.57 | 0.02 |
| | CG | CR | -1.33 | 0.57 | 0.02 |
| | | FR | -1.36 | 0.57 | 0.02 |
| Translation | CR | FR | 0.24 | 0.45 | 0.59 |
| | | CG | 0.77 | 0.45 | 0.09 |
| | FR | CR | -0.24 | 0.45 | 0.59 |
| | | CG | 0.53 | 0.45 | 0.24 |
| | CG | CR | -0.77 | 0.45 | 0.09 |
| | | FR | -0.53 | 0.45 | 0.24 |
| Writing | CR | FR | -0.44 | 0.25 | 0.08 |
| | | CG | 0.08 | 0.25 | 0.76 |
| | FR | CR | 0.44 | 0.25 | 0.08 |
| | | CG | 0.52 | 0.25 | 0.04 |
| | CG | CR | -0.08 | 0.25 | 0.76 |
| | | FR | -0.52 | 0.25 | 0.04 |
| Total | CR | FR | -1.05 | 1.87 | 0.58 |
| | | CG | 4.00 | 1.87 | 0.04 |
| | FR | CR | 1.05 | 1.87 | 0.58 |
| | | CG | 5.05 | 1.87 | 0.01 |
| | CG | CR | -4.00 | 1.87 | 0.04 |
| | | FR | -5.05 | 1.87 | 0.01 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Given that the mid-test served as a monitor test showing the how extensive reading worked half way through the present study and the second half of the program was to begin, what caused the significant differences between the experimental and control groups and the differences between the two experimental groups remained to be examined in the post-test. Whether the statistically significant differences were going to be larger or the statistically insignificant difference would become significant in the post-test was unknown by the mid-test.

At this stage, the explanation was that the Complementary Reading group showed its advantage in reading mainly because of their huge reading amount. According to Table 4.2, by the mid-test the Complementary Reading group read 417,864 words on average, nearly four times as much as that of the Free Reading group. Simultaneously, the complementary reading treatment involved devoting a considerable proportion of class time to in-class discussion (approximately ten minutes, see Section 3.6.2 for detailed introduction of the complementary reading treatment), which accordingly reduced the time spent on in-class form-focused teaching. The researcher assigned less homework to the Complementary Reading group to guarantee their home reading time. It was probably another reason for their lower marks than the Free Reading group in grammar. Another possibility that may account for the difference was that in language performance, listening and reading belong to the respective skill category, while speaking and writing are referred to as productive skills. We were waiting to see whether it would take time for the long-term benefits of extensive reading on productive skills to emerge, or extensive reading has less to offer to productive skills than the Grammar-Translation Method. It was well worth anticipating whether the gaps would be narrowed or widened in the post-test, which may further explain the differences.

The post-test was administered in June, 2011, which was also the end of year English test (see Section 3.3.3 for the detailed information of the post-test). The situation changed a lot compared with that of the mid-test. As shown in Table 4.12 below, the Free Reading group has made advantage in terms of mean score in all proficiency areas and language knowledge components covered in the

post-test. There was a significant main effect for the treatments as a factor in listening, grammar, cloze, reading, writing and total score. It is obviously visible in the post-test that there was a significant main effect for the treatments as a factor in every proficiency area and language knowledge component tested except vocabulary.

Table 4. 12 Descriptive statistics and results of ANOVA from the post-test

| Section (Total score) | Group | N | Mean | SD | F | Sig. |
|-----------------------|-------|----|-------|------|-------|------|
| Listening (20) | CR | 33 | 14.48 | 1.77 | 3.38 | 0.04 |
| | FR | 33 | 15.42 | 2.31 | | |
| | CG | 33 | 14.06 | 2.41 | | |
| Grammar (20) | CR | 33 | 15.88 | 2.67 | 2.68 | 0.07 |
| | FR | 33 | 16.64 | 2.22 | | |
| | CG | 33 | 15.03 | 3.44 | | |
| Vocabulary (9) | CR | 33 | 7.15 | 1.39 | 1.92 | 0.15 |
| | FR | 33 | 7.39 | 1.62 | | |
| | CG | 33 | 6.61 | 1.95 | | |
| Cloze (10) | CR | 33 | 6.55 | 1.35 | 9.20 | 0.00 |
| | FR | 33 | 6.79 | 1.63 | | |
| | CG | 33 | 5.33 | 1.43 | | |
| Reading (16) | CR | 33 | 11.55 | 1.95 | 5.33 | 0.01 |
| | FR | 33 | 11.58 | 2.08 | | |
| | CG | 33 | 10.24 | 1.62 | | |
| Translation (15) | CR | 33 | 10.55 | 1.42 | 6.48 | 0.00 |
| | FR | 33 | 11.76 | 1.40 | | |
| | CG | 33 | 11.15 | 1.28 | | |
| Writing (10) | CR | 33 | 5.57 | 0.73 | 8.19 | 0.00 |
| | FR | 33 | 6.39 | 0.77 | | |
| | CG | 33 | 6.07 | 0.98 | | |
| Total (100) | CR | 33 | 71.72 | 5.67 | 10.21 | 0.00 |
| | FR | 33 | 75.97 | 6.57 | | |
| | CG | 33 | 68.50 | 7.81 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Table 4.13 below confirms that the variance of the scores for the three groups are roughly equal before we go further with the results of ANOVA presented in Table 4.12 above.

Table 4. 13 Test of homogeneity of variances from the post-test

| Section (Total score) | Levene Statistic | df1 | df2 | Sig. |
|-----------------------|------------------|-----|-----|------|
| Listening (20) | 1.158 | 2 | 96 | .319 |
| Grammar (20) | 2.326 | 2 | 96 | .103 |
| Vocabulary (9) | 2.280 | 2 | 96 | .108 |
| Cloze (10) | .769 | 2 | 96 | .466 |
| Reading (16) | 1.009 | 2 | 96 | .368 |
| Translation (15) | .481 | 2 | 96 | .619 |
| Writing (10) | 1.308 | 2 | 96 | .275 |
| Total (100) | 1.247 | 2 | 96 | .292 |

* The mean difference is significant at the 0.05 level.

Table 4. 14 The results of multiple comparisons by the Fisher's LSD test in the post-test

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| Listening | CR | FR | -0.94 | 0.54 | 0.08 |
| | | CG | 0.42 | 0.54 | 0.43 |
| | FR | CR | 0.94 | 0.54 | 0.08 |
| | | CG | 1.36 | 0.54 | 0.01 |
| | CG | CR | -0.42 | 0.54 | 0.43 |
| | | FR | -1.36 | 0.54 | 0.01 |
| Grammar | CR | FR | -0.76 | 0.69 | 0.28 |
| | | CG | 0.85 | 0.69 | 0.22 |
| | FR | CR | 0.76 | 0.69 | 0.28 |
| | | CG | 1.61 | 0.69 | 0.02 |
| | CG | CR | -0.85 | 0.69 | 0.22 |
| | | FR | -1.61 | 0.69 | 0.02 |
| Vocabulary | CR | FR | -0.24 | 0.41 | 0.56 |
| | | CG | 0.55 | 0.41 | 0.19 |
| | FR | CR | 0.24 | 0.41 | 0.56 |
| | | CG | 0.79 | 0.41 | 0.06 |
| | CG | CR | -0.55 | 0.41 | 0.19 |
| | | FR | -0.79 | 0.41 | 0.06 |
| Cloze | CR | FR | -0.24 | 0.36 | 0.51 |
| | | CG | 1.21 | 0.36 | 0.00 |
| | FR | CR | 0.24 | 0.36 | 0.51 |
| | | CG | 1.45 | 0.36 | 0.00 |
| | CG | CR | -1.21 | 0.36 | 0.00 |
| | | FR | -1.45 | 0.36 | 0.00 |
| Reading | CR | FR | -0.03 | 0.47 | 0.95 |
| | | CG | 1.30 | 0.47 | 0.01 |
| | FR | CR | 0.03 | 0.47 | 0.95 |
| | | CG | 1.33 | 0.47 | 0.01 |
| | CG | CR | -1.30 | 0.47 | 0.01 |
| | | FR | -1.33 | 0.47 | 0.01 |
| Translation | CR | FR | -1.21 | 0.34 | 0.00 |
| | | CG | -0.61 | 0.34 | 0.07 |
| | FR | CR | 1.21 | 0.34 | 0.00 |
| | | CG | 0.61 | 0.34 | 0.07 |
| | CG | CR | 0.61 | 0.34 | 0.07 |
| | | FR | -0.61 | 0.34 | 0.07 |
| Writing | CR | FR | -0.82 | 0.21 | 0.00 |
| | | CG | -0.50 | 0.21 | 0.02 |
| | FR | CR | 0.82 | 0.21 | 0.00 |
| | | CG | 0.32 | 0.21 | 0.12 |
| | CG | CR | 0.50 | 0.21 | 0.02 |
| | | FR | -0.32 | 0.21 | 0.12 |
| Total | CR | FR | -4.25 | 1.66 | 0.01 |
| | | CG | 3.22 | 1.66 | 0.05 |
| | FR | CR | 4.25 | 1.66 | 0.01 |
| | | CG | 7.47 | 1.66 | 0.00 |
| | CG | CR | -3.22 | 1.66 | 0.05 |
| | | FR | -7.47 | 1.66 | 0.00 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Table 4.14 above shows where the differences are as detected by the post hoc Fisher's LSD test. The two experimental groups were significantly better than the

control group in cloze, reading and total score. It did not mean the two experimental groups were equally good. In comparison, the Free Reading group was significantly better at translation and writing than the Complementary Reading group, and the former's total score was slightly better than the latter's. More surprisingly, as the group who read most, the Complementary Reading group was weaker than the Control Group in translation and writing even though its total score was significant higher than that of the Control Group ($p=0.05$). What caused the differences between the pre- and post-test will be further discussed in subsequent discussion sections.

Table 4. 15 Descriptive statistics and results of ANOVA from the delayed post-test

| Section (Total score) | Group | N | Mean | SD | Std. Error | F | Sig. |
|-----------------------|-----------------------|----|--------|-------|------------|------|------|
| Listening (30) | CR | 32 | 22.38 | 2.21 | 0.39 | 1.93 | 0.15 |
| | FR | 33 | 22.67 | 2.78 | 0.48 | | |
| | CG | 32 | 21.41 | 3.04 | 0.54 | | |
| Grammar (16) | CR | 32 | 12.41 | 2.00 | 0.35 | 1.89 | 0.16 |
| | FR | 33 | 12.30 | 1.94 | 0.34 | | |
| | CG | 32 | 11.53 | 1.97 | 0.35 | | |
| Vocabulary (9) | CR | 32 | 6.19 | 1.51 | 0.27 | 0.00 | 1.00 |
| | FR | 33 | 6.21 | 1.11 | 0.19 | | |
| | CG | 32 | 6.22 | 1.39 | 0.24 | | |
| Cloze (15) | CR | 32 | 11.31 | 1.79 | 0.32 | 2.29 | 0.11 |
| | FR | 33 | 10.97 | 1.67 | 0.29 | | |
| | CG | 32 | 10.38 | 1.86 | 0.33 | | |
| Reading 1(27) | CR | 32 | 21.28 | 2.68 | 0.47 | 0.20 | 0.82 |
| | FR | 33 | 21.09 | 3.47 | 0.60 | | |
| | CG | 32 | 20.75 | 3.88 | 0.69 | | |
| Reading 2 (8) | CR | 32 | 5.20 | 1.65 | 0.29 | 2.06 | 0.13 |
| | FR | 33 | 5.88 | 1.63 | 0.28 | | |
| | CG | 32 | 5.19 | 1.42 | 0.25 | | |
| Translation (20) | CR | 32 | 14.11 | 2.05 | 0.36 | 4.65 | 0.01 |
| | FR | 33 | 14.41 | 1.83 | 0.32 | | |
| | CG | 32 | 13.04 | 1.78 | 0.31 | | |
| Writing (25) | Content (10) | CR | 32 | 5.86 | 0.64 | 5.27 | 0.01 |
| | | FR | 33 | 6.23 | 0.63 | | |
| | | CG | 32 | 5.70 | 0.74 | | |
| | Language quality (10) | CR | 32 | 5.70 | 0.66 | 4.95 | 0.01 |
| | | FR | 33 | 6.08 | 0.70 | | |
| | | CG | 32 | 5.53 | 0.78 | | |
| | Organization (5) | CR | 32 | 1.81 | 0.45 | 0.22 | 0.80 |
| | | FR | 33 | 1.89 | 0.45 | | |
| | | CG | 32 | 1.84 | 0.57 | | |
| | Total (25) | CR | 32 | 13.38 | 1.40 | 4.35 | 0.02 |
| | | FR | 33 | 14.20 | 1.48 | | |
| | | CG | 32 | 13.08 | 1.85 | | |
| Total (150) | CR | 32 | 106.25 | 8.05 | 1.42 | 4.41 | 0.01 |
| | FR | 33 | 107.72 | 8.43 | 1.47 | | |
| | CG | 32 | 101.59 | 9.45 | 1.67 | | |

* The mean difference is significant at the 0.05 level.

Note 1: The results of Levene tests show there is no deviation from homogeneity of variance.

Note 2: Reading 1 was in the form of multiple choice and Reading 2 was short questions and answers.

The delayed post-test was administered in June 2012, one year after the program ended (see Section 3.3.4 for the detailed information of the delayed post-test). A note to make here is that one participant in the Complementary Reading group and one participant in the Control group did not take the delayed-test because of physical discomfort. The situation differed from that of the post-test. As shown in Table 4.15 above, there was a significant main effect for the treatments as a factor in translation, writing and total score. In particular, there was a significant main effect for the treatments as a factor in content and language quality in the writing section.

Table 4. 16 Test of homogeneity of variances from the delayed post-test

| Section (Total score) | | Levene Statistic | df1 | df2 | Sig. |
|-----------------------|-----------------------|------------------|-----|-----|------|
| Listening (30) | | 1.79 | 2 | 94 | 0.17 |
| Grammar (16) | | 0.12 | 2 | 94 | 0.88 |
| Vocabulary (9) | | 1.58 | 2 | 94 | 0.21 |
| Cloze (15) | | 0.45 | 2 | 94 | 0.64 |
| Reading 1 (27) | | 2.00 | 2 | 94 | 0.14 |
| Reading 2 (8) | | 0.25 | 2 | 94 | 0.78 |
| Translation (20) | | 0.09 | 2 | 94 | 0.92 |
| Writing (25) | Content (10) | 0.31 | 2 | 94 | 0.73 |
| | Language quality (10) | 0.51 | 2 | 94 | 0.60 |
| | Organization (5) | 1.55 | 2 | 94 | 0.22 |
| | Total (25) | 0.96 | 2 | 94 | 0.39 |
| Total (150) | | 0.43 | 2 | 94 | 0.65 |

* The mean difference is significant at the 0.05 level.

Note: Reading 1 was in the form of multiple choice and Reading 2 was short questions and answers.

4.3.2 The EPER Placement Test B

The EPER Placement Test B (see Section 3.3.4 for the detailed information of the EPER Placement Test B) was administered to the three groups in September, 2010, the beginning of the project. In June, 2011, the same test was re-administered to the three groups. Table 4.17 below shows results of the pre- and post-EPER Placement tests.

Table 4. 17 Results of the pre- and post-EPER Placement Test B

| EPER Level | | A | B | C | D | E | Total |
|------------|--------|------|-------|-------|-------|------|--------|
| Headwords | | 2200 | 1900 | 1600 | 1200 | 800 | / |
| Pre-test | Number | 1 | 19 | 41 | 36 | 2 | 99 |
| | % | 1.01 | 19.19 | 41.41 | 36.36 | 2.02 | 100.00 |
| Post-test | Number | 5 | 40 | 37 | 17 | / | 99 |
| | % | 5.05 | 40.40 | 37.37 | 17.17 | / | 100.00 |

The comparison between the pre- and post-EPER tests showed the change in EPER levels of the whole experimental population. As shown in Table 4.17, the absolute numbers of Level A and B participants increases sharply. The percentage of Level A participants increased from 1% to 5%, five times as many as that in the pre-test. The number of Level B participants doubled, from 19% to 40%. In addition, the number of Level D participants decreased by 50%, and there was no Level E participant. A general trend was observed from Table 4.17 that the overall EPER levels of the whole experimental population move onto a higher stage.

Table 4.18, the Group*Progress cross tabulation, illustrates that ten students in the whole experimental population slipped back by one level. Thirty-eight participants remained stable in terms of their EPER levels. Forty-three participants, the largest group in the cross tabulation, moved forward by one level. Seven participants made progress by two levels and one participant advanced by three levels.

Table 4. 18 Group * Progress Cross tabulation of the EPER Placement Test of the whole sample

| Progress in level | | -1 | 0 | 1 | 2 | 3 | Total |
|-------------------|-------|----|----|----|---|---|-------|
| Group | CR | 1 | 10 | 16 | 5 | 1 | 33 |
| | FR | 4 | 11 | 16 | 2 | 0 | 33 |
| | CG | 5 | 17 | 11 | 0 | 0 | 33 |
| | Total | 10 | 38 | 43 | 7 | 1 | 99 |

The Chi-square test of the Group * Progress Cross tabulation presented in Table 4.19 below confirmed that the EPER Placement Test produced evidence of statistically significant differences between pre- and post-tests in terms of the participants' EPER levels ($p < 0.01$, Chi-square=18.26, df=4). The progress in EPER levels did not happen by chance or sample errors but was attributed to extensive reading.

In this section, the EPER Placement test results of the whole experimental population have been presented. The difference within individual groups in the EPER Placement Test B will be presented and discussed in details in Section 4.4.6 Extensive reading and EPER levels.

Table 4. 19 Chi-Square Tests of the results of the pre- and post-EPER Placement Test

| Chi-Square Tests | | | |
|--------------------|------------|-------|-----------------------|
| | Value | df | Asymp. Sig. (2-sided) |
| Pearson Chi-Square | 18.26 | 4 | 0.001 |
| Likelihood Ratio | 19.55 | 4 | 0.001 |
| N of Valid Cases | 198 | | |
| Symmetric Measures | | | |
| | | Value | Approx. Sig. |
| Nominal by Nominal | Phi | .304 | .001 |
| | Cramer's V | .304 | .001 |
| N of Valid Cases | | 198 | |

* The mean difference is significant at the 0.05 level.

4.4 Extensive reading and language knowledge

Prior to interpreting and discussing the results of the pre-, mid- and post-tests and the EPER Placement Test B, it must be noted here that the subsequent interpretations and discussions are to be classified into three categories and presented in the following three corresponding sections: (1) extensive reading and language knowledge, i.e., vocabulary and grammar in Section 4.4, (2) extensive reading and receptive skills, i.e., reading and listening in Section 4.5, and (3) extensive reading and productive skills, i.e., translation and writing in Section 4.6. As you may be aware, this order differs from that in the previous section, in which the test results are presented in light of the subsections' availability and sequences in the test papers. This modification of the order of the subsections is intended to reflect a normal expectation that language knowledge components and particular skills directly related to reading be carried out first.

4.4.1 Extensive reading and vocabulary acquisition

Words are so pervasive in our life and so central to language that we do not often stop to reflect on teaching and learning of vocabulary in language pedagogy. Theoretical priorities on vocabulary have changed throughout language teaching history, as reflected in the relative importance placed on it in terms of grammatical function, meaning, and collocation or conversion. To place vocabulary teaching and learning in its proper perspective, Nation (2001) summarizes the following four major strands in language classroom and points out that balanced vocabulary instruction should consist of the following four major strands.

The first strand is learning from comprehensible meaning-focused input, which is closely associated with incidental vocabulary acquisition, either acquiring new lexical items or expanding knowledge of lexical items that have been known to the readers, without focusing on form (e.g., Schmitt, 2000; Stahl & Nagy, 2006). For successful learning from comprehensible meaning-focused input to occur, a mastery of at least 95% of the running words is the premise for successful reading with instructional support, and a mastery of 99% of the running words in the text is the expectation for independent fluent reading (Nation, 2001). The second strand is form-focused instruction, i.e., the direct teaching and learning of lexical items. The third strand is meaning-focused output, through which learners expand their vocabulary knowledge in speaking and writing activities. The fourth strand in a balanced course is fluency development, which does not aim to teach new lexical items to learners but enable them to be more fluent in using lexical items they have already known. To summarize, the first two strands are receptive and the last two are productive.

How well are these four strands incorporated into vocabulary teaching and learning in Shanghai senior high schools? Prior to encountering it incidentally in reading, the students will encounter a new word in a bilingual list accompanied by its Chinese equivalents or translations. If considered a core word, it will be elaborated in classroom instruction by the teacher. If it is a noun, for example *curiosity*, the teachers will provide some phrases containing it such as *out of curiosity*, *curiosity killed the cat*, to help the students explore its various meanings and remember it. If it is a verb, for example *occur*, the teacher will show the students its collocations and sentence patterns such as “*it occurs to sb that...*” and explain its spelling and grammatical rules in case the students should make mistakes when using it in their translation and writing. If it is an adjective, for example *heavy*, the teacher will provide its usual collocation as chunks, for example *a heavy rain*. In general, the teachers will explain all the usage related to a particular word as comprehensively as they can in light of their own teaching experience and assessment requirements. The students are required to make careful notes of what is taught by their teachers in case they should neglect or forget some important issues that need to be taken care of when using it. One or two days later after the word has been taught, the teachers will give the students a

dictation to check whether the students have mastered its spelling and meaning. Later, the teachers will design some grammar and translation exercise to help the students recycle the words and put it into productive use. But due to the time constraints in class, no more attention can be allocated to it since the teachers need to move to the next chapter and work with other new words.

It is obvious that the second and the third strands in Nation's framework, the language-focused or form-focused instruction and meaning-focused output, are well incorporated into in-class vocabulary instruction, while the other two fail to be clearly reflected. From the students' perspective, they acquire their vocabulary in a passive manner. Firstly, they are not able to choose which word to learn. Secondly, they learn a so-called "core" word intensively just because their teachers tell them it is important, but they may not understand where its importance lies. Thirdly, a new word is usually presented in a sentence that elucidates its meaning in the text, but chances that they encounter it incidentally in extra reading is low because the time they spend on reading in English is rather limited. When asked about vocabulary instruction in school, the students probably remember reciting bilingual lists of words with dictations every day and tests at regular intervals. Last but not least, in the whole process the students are not provided with the chance to recycle the words orally due to the negative backwash effect that the NMET brings. Since speaking is not tested in the NMET, most of the teachers tend to focus their attention on semantic and grammatical functions of the words rather than pronunciation and oral productions. Rarely do the students have the chance to practice pronunciation in the classroom or are they able to communicate with their teachers or classmates orally using the newly-taught words. As discussed in the section extensive reading and listening proficiency, students having difficulty pronouncing the words accurately are more likely to encounter difficulty in recognizing the phonological information in listening, which in turn impedes their receiving linguistic input in oral form and eventually affects their second language acquisition. From the teachers' perspective, they also complain that intensive instruction for every word is not feasible. How can they carry out vocabulary instruction in great breadth and depth and still have time left for anything else in the curriculum?

As can be seen, the most obvious drawback of the current practice of vocabulary instruction in Shanghai senior high schools is the lack of comprehensible meaning-focused input, a core component for incidental vocabulary acquisition and fluency development. To present a balanced vocabulary instruction, a component of comprehensible meaning-focused input must be incorporated to facilitate vocabulary acquisition and fluency development.

The existing body of literature has well documented the significance of extensive reading in learners' vocabulary development. As discussed in Section 2.1.1, there are three styles of defining extensive reading: extensive reading as a reading style, extensive reading as an input medium and extensive reading as a teaching approach. The value of defining extensive reading as a style is that extensive reading can be associated with incidental vocabulary acquisition and fluency development (see Section 2.1.1 for detailed discussion).

The significance of extensive reading on incidental vocabulary acquisition and fluency development is well supported by the research evidence from empirical studies. The empirical studies first establish that new lexical items can be acquired incidentally from extensive reading (e.g., Brown et al., 2008). Based on this, the general tendency observed is that the more frequently the items appear in the text, the more likely they are to be acquired (e.g., Kweon & Kim, 2008). The repeated exposure to frequent words can be realized by using reading materials with a common theme or written by the same author (e.g., Gardner, 2008). The limitation of the repeated exposure of lexical items through extensive reading is that extensive reading leads to better acquisition of word meaning, whereas its effect on word use is relatively moderate (Alshwaikh, 2004).

In addition to the studies noted above, there are many similar studies supporting the effect of extensive reading on incidental vocabulary acquisition (e.g., Cho & Krashen, 1994; Coday, 1997; Coady & Huckin, 1997; Day, et al., 1991; Horst, 2000; Horest, et al., 1998; Huckin & Coady, 1999; Paribakht & Wesche, 1997; Poulshock, 2010; Prichard, 2008; Pudilo & Hambrick, 2008; Rosszell, 2007; Sonbul & Schmitt, 2010), and due to the limited space we just note a few here. These studies indicate the following assumptions which may explain the

relationship between extensive reading and vocabulary acquisition: the “noticing” assumption, the “guessing ability” assumption, the “guessing-retention link” assumption, and the “cumulative gain” assumption. These assumptions are ordered in a clear logic: readers are exposed to the target word through repeated exposure, so they notice it. The word is important for understanding the text, so they guess its meaning from the context. They have been exposed to the target word several times, and thus they try to guess from their retention. They know its meaning at the first time, and they know its part of speech the next time they meet with it.

The above assumptions are supported by empirical studies. In Brown et al.’s study (2008), the more frequently the items appear in the text, the more likely they are to be acquired by learners, which supports the “noticing” assumption in a sense that the retention of words is closely related to the words’ frequency. Gardner (2008) further strengthens this assumption by using collections of reading materials with a common theme and authorship. In Alshwaikh’s (2004) study, the readers scored higher on word meaning levels than the non-readers, which showed that the acquisition of words from reading were cumulative.

The major challenge to the above assumptions comes from Laufer (2003). In a series of three studies, she provided different conditions for vocabulary learning. In the first study, sixty Arabic university students who learned English as a foreign language were stratified into two groups. The students in the first group were requested to do a reading comprehension task. The ten target words in the reading text were glossed in Hebrew, the first language of the students. The students in the second group were required to produce sentences with the same ten target words. In the second study, the eighty-two EFL university students were also stratified into two groups. The first group repeated the same reading comprehension task in the first study. The second group was required to write an essay using the same ten target words. In the third study, ninety high school students were stratified into three groups, unlike the two-group design in the previous two studies. The first group repeated the reading comprehension task. They read a text containing the same ten target words which were not glossed, and they were allowed to refer to a dictionary when necessary. The second group

was asked to produce sentences with the same target words. The translations and parts of speech of the ten target words were written on a piece of paper, and the paper was made available to the participants. The task for the third group was sentence completion. The participants in this group would need to complete the ten sentences by using the same ten target words, one word for each sentence. They were allowed to consult a bilingual dictionary. Despite the differences in research design, the three studies shared the same underlying hypothesis that vocabulary learning through “conscious” tasks would be greater than through reading. This hypothesis was supported by the consistent findings from the three studies, showing that engaging with the words more “consciously” produced better gains in vocabulary learning than through reading.

The results of Laufer’s study seem to be conflicting with the above-noted ones, but in fact they share a lot in common. Firstly, it is assumed that comprehensible meaning-focused input engages learners in incidental vocabulary acquisition, because they are immersed in an input-rich context and exposed to abundant linguistic items and structures. Through repeated exposure, the readers notice some items, and learn them. In their subsequent meeting with the words, they know the words better and cumulate more word knowledge. In the previous studies, the readers “incidentally” acquire the words, and in Laufer’s study, the readers learn the words “consciously” through tasks. From the surface the participants in these different studies acquire the words in different ways, but in nature in the same way: notice them, and then learn them. The only difference between the studies is that Laufer told the participants which words to concentrate on, while in other studies the participants identify the key words by frequency.

In our study, it seems reasonable to expect remarkable gains in terms of vocabulary knowledge and size at the end of the present study. Surprisingly, there is a significant main effect for the treatments as a factor in the vocabulary section in the mid-test ($p=0.01$) instead of the post-test as shown in Table 4.20 below.

Table 4. 20 Descriptive statistics and results of ANOVA from the vocabulary sections in the mid- and post-tests

| Test (Total score) | Group | N | Mean | SD | F | Sig. |
|--------------------|-------|----|------|------|------|------|
| Mid-test (9) | CR | 33 | 8.33 | 1.19 | 5.46 | 0.01 |
| | FR | 33 | 8.18 | 1.18 | | |
| | CG | 33 | 7.36 | 1.45 | | |
| Post-test (9) | CR | 33 | 7.15 | 1.39 | 1.92 | 0.15 |
| | FR | 33 | 7.39 | 1.62 | | |
| | CG | 33 | 6.61 | 1.95 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

The multiple comparisons by the Fisher's LSD test in the mid-test among the three groups clearly illustrated that there was a statistically significant difference between the experimental groups and the control group (see Table 4.21 below). The explanation for this result at this stage is that the two experimental groups did extensive reading while the Control groups did not. Further, the two experimental groups were expected to outperform the Control group in the post-test as well and there should be a significant main effect for the treatments as a factor in vocabulary, because by the post-test the two experimental groups would have read much more than they did by the mid-test.

Table 4. 21 The results of multiple comparisons by the Fisher's LSD test in the vocabulary sections in the mid- and post-tests

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| Mid-test | CR | FR | 0.15 | 0.32 | 0.63 |
| | | CG | 0.97 | 0.32 | 0.00 |
| | FR | CR | -0.15 | 0.32 | 0.63 |
| | | CG | 0.82 | 0.32 | 0.01 |
| | CG | CR | -0.97 | 0.32 | 0.00 |
| | | FR | -0.82 | 0.32 | 0.01 |
| Post-test | CR | FR | -0.24 | 0.41 | 0.56 |
| | | CG | 0.55 | 0.41 | 0.19 |
| | FR | CR | 0.24 | 0.41 | 0.56 |
| | | CG | 0.79 | 0.41 | 0.06 |
| | CG | CR | -0.55 | 0.41 | 0.19 |
| | | FR | -0.79 | 0.41 | 0.06 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

However, there is no significant main effect for the treatments as a factor in the vocabulary section in the post test ($p=0.15$) (see Table 4.21 above). The multiple comparisons by the Fisher's LSD test in the post-test between the three groups showed that there was no statistically significant difference between each two

groups, although the two experimental groups still obtained advantage in terms of mean score (CR=7.15. FR=7.39, CG=6.61).

Our next step is to conduct a fine-grained analysis of the three groups' performance on each item in the vocabulary section in the post-test in order to gain deeper insights into the relationship between extensive reading and vocabulary acquisition. The usual way to analyze the lexical items that are tested in a proficiency test is to make reference to the published word lists that indicate the frequency with which the words have been found to be used (Heaton, 1975; Read, 2000). In our discussion, the General Service List (GSL), the Word Frequency Book (WFB) and the Academic Word List (AWL) are chosen as classic reference lists. Given that these three lists were published at least more than ten years ago although they are landmarks in lexical studies, the Corpus of Contemporary American English (1990-present) (COCA) serve as the most recent language data for comparison.

The General Service List (West, 1953) enlists two thousand entries which are considered to be of “general service” value to users of English. This does not mean only two thousand single words are in active use or of grate service value to the learners. In the GSL, one includes words from the same word family. For example, under the entry “ABLE”, there is also “ability”, and their meanings and frequencies. In this sense, an entry is a headword representing a word family (see Section 3.4 for introduction of headword and word family). The greatest pedagogical application of the GSL is that it used to serve as base word list for developing materials for learners of English (not native speakers), for example, graded readers (see Section 2.3 for lexical control in graded readers). The Word Frequency Book (Carroll et al., 1971) presents the 86,741 different word types found in the 5,088,721 words of running text which secondary students in America are exposed to, which makes it meaningful to use WFB as a reference word list when the test takers of the post-test come from secondary schools in China. The Academic Word List (Coxhead, 1998) includes 570 high-frequency headwords from academic texts written in English. Unlike GSL and WFB, in which words are presented in alphabetic order, AWL divides the 570 words into ten sub-lists according to their frequency. For instance, the sixty most frequent

words appear in the first sub-list, and the next most frequent sixty words are in the second sub-list, and so on. That is to say, the latter sub-list a word appears in, the less frequent it is in academic texts. BNC is a corpus of British English which contains 100 million words divided among a comprehensive range of spoken and written texts (available at: <http://www.natcorp.ox.ac.uk/>). COCA is a corpus of American English, with a 450-million words collection of samples from a wide range of discourses and genres, including spoken, newspapers and academic texts (available at: <http://corpus.byu.edu/coca/>). The 450 million words are accumulated by 20 million each from the year 1990 to 2012, with the most recent texts from summer, 2012).

In addition to the reference lists and corpora for examining the frequency of the ten words, the *Range* program designed by Paul Nation and Alex Hartley will also be run. *Range* has a number of useful functions (please visit: <http://www.victoria.ac.nz/lals/about/staff/paul-nation>, for detailed information and downloading). For the present study, *Range* will provide information on the ten words' frequency level in BNC. The designers of *Range* have divided the word list of BNC into twenty-five sub-lists. Sub-list 1 contains the most frequent 1,000 words in BNC, and Sub-list 2 contains the 2nd most frequent 1,000 words in BNC, and so on. *Range* will tell us at which frequency level the ten words are.

The ten words tested in the vocabulary section in the post-test were: *find*, *traditional*, *disturb*, *restore*, *useless*, *performance*, *carefully*, *hesitate*, *suddenly* and *show* (see Appendix C). As shown in Table 4.22 below, nine of the ten words tested in the vocabulary section in the post-test have the following characteristics. First, the ten words tested can be classified into three categories according to their syntactical functions: noun, verb and adjective, all of which are content words (see Section 3.4 for the discussion on content words and function words). Words in the same category are similar neither in meanings nor usages. Second, all of the ten words appear in WFB, seven are found in GSL and only two are available in AWL. Three of the ten words, *restore*, *hesitate* and *disturb* were explicitly taught in class because they were in the vocabulary list attached to the textbooks. According to their frequency and availability in WFB, GSL and AWL, *restore*, *hesitate* and *disturb* are not high-frequency words or high-frequency

academic words, nor are they of great “general service” to learners of English. Interestingly, findings from *Range* are consistent with the findings from the three reference lists. The ten words are distributed to three frequency levels in BNC: *find*, *show*, *suddenly*, *carefully* and *useless* are at frequency Level 1, i.e., they are in the most frequent one thousand words in BNC, *performance* and *traditional* are at frequency Level 2, i.e., they are in the second most frequent one thousand words in BNC, and *restore*, *disturb* and *hesitate* at frequency Level 3, i.e., they are in the third most frequent one thousand words in BNC. It is the three words at the third frequency level in BNC that were explicitly taught in class. Obviously, the selection of the ten words constitutes a representative sample of two types of words: high-frequency words in texts written in English and words which were explicitly taught in class, although they are not of great service or academic value. This constitution, on one hand, assesses how the test-takers manage the flow of the information by placing the high-frequency words in the right place; on the other hand, tests whether they are able to handle words that have been explicitly taught in class.

Within the same frequency level, a detailed analysis of the words across COCA, WFB and GSL shows that the two most frequent words are *show* and *find*. As the more frequent one in these two words, *find* appears in WFB 11,758 times in all, with the base form *find/Find* 6,921 times, and inflectional forms *finds* 1,079 times, *found* 3,362 times and *finding* 396 times. Find is also the most frequent words of the ten according to the GSL, which appeared 3,874 times. In comparison, *find* is also the most frequent word in the ten in COCA. It appears 463,482 times in all, with the base form *find/Find* 19,061 times, and the inflectional forms *finds* 195,160 times, *found* 213,390 times and *finding* 35,871 times.

According to WFB and GSL, *show* is the second most frequent words in the ten. *Show* appears in WFB 4,751 times in total, with the base form *show* 2,734 times, *Show* 24 times, and the inflectional forms *showed* 478 times, *shows* 1,184 times and *showing* 331 times. It appears in GSL 3,192 times. Consistent with the findings from WFB and GSL, *show* is also the second most frequent one in the

Table 4. 22 Detailed analysis of the ten words tested in the vocabulary section in the post-test

| Words | Syntactical functions | Frequency level in BNC | Frequency in COCA | Frequency in WFB | Frequency in GSL | Frequency Level in AWL | Explicitly taught in class? |
|-------------|-----------------------|------------------------|-------------------|------------------|------------------|------------------------|-----------------------------|
| find | v. | 1 | 19,061/463,482 | 6,921/11,758 | 3,874 | (found) 9 | No. |
| show | v./n. | 1 | 175,093/344,710 | 2,734/4,751 | 3,192 | N/A | No. |
| suddenly | adv. | 1 | 48,442 | 993 | 554 | N/A | No. |
| carefully | adv. | 1 | 25,835 | 993 | 244 | N/A | No. |
| useless | adj. | 1 | 4,562 | 86 | 82 | N/A | No. |
| performance | n. | 2 | 60,944 | 149 | 344 | N/A | No. |
| traditional | adj. | 2 | 49,185 | 90 | N/A | N/A | No. |
| restore | v. | 3 | 7,776/16,944 | 34/82 | N/A | (found) 8 | Yes. |
| disturb | v. | 3 | 1,800/13,297 | 22/100 | 177 | N/A | Yes. |
| hesitate | v. | 3 | 2,542/8,601 | 29/98 | 116 | N/A | Yes. |

ten, which appears in 344,710 times in total, with the base form *show/Show* 175,093 times, *showed* 47,289 times, *shows* 59,920 times, *shown* 36,095 and *showing* 26,313 times.

As the least frequent words in the ten in WFB, *restore* appears in WFB 82 times in all, with the base form *restore* 34 times, and the inflectional forms *restored* 39 times, *restores* 2 times and *restoring* 7 times. Unlike its frequency in WFB, *restore* becomes the most frequent one in the three 3rd frequency level words. It appears in COCA 16,944 times in all, with the base form *restore/Restore* 7,776 times, and the inflectional forms *restored* 5,871 times, *restores* 490 times, and *restoring* 2,807 times. The reason that may explain the discrepancy between a classic reference list published in the 1970s and one of the most recent English language corpus is that *restore* is found to be an academic words which appears in the 8th sub-list in AWL. Its academic property may expand its use in at least one more genre, and thus increases its frequency in COCA.

Unlike their frequencies in the classic reference lists and the most updated English language corpus, the ten words had different frequencies in the textbooks and received different levels of instruction or teachers' concern in teaching. Among them, *restore*, *disturb* and *hesitate* appeared in Senior One textbooks, and were explicitly taught by teachers.

The above comparison reveals a discrepancy in the ten words' frequency and importance in corpora and daily teaching. It would be interesting to see how the three groups of participants performed on these ten words exactly. Table 4.23 below illustrates the detailed statistic of the three groups' performance on the nine words. In the post-test, there were nine blanks in the passage, and ten words were provided as options. According to the key, *restore* was not used. So in Table 4.23 below, statistics of nine words are presented.

Two interesting findings emerge from Table 4.23. First, there is a significant main effect for the treatments as a factor on Item 42, namely *show*, which is a high-frequency word both in the classic reference lists and contemporary English language corpora. In addition, a practical difference between the three groups is

found on Item 46, namely *find*, which is also a high frequency word in both classic reference lists and contemporary corpora, although the difference is not statistically significant (CR=0.94, FR=0.91, CG=0.76, $p=0.06$).

Table 4. 23 Descriptive statistics and results of ANOVA from the vocabulary section in the post-test

| Item | Group | N | Mean | SD | Std. Error | F | Sig. |
|---------------------|-------|----|------|------|------------|------|------|
| 41 (useless) | CR | 33 | 0.61 | 0.50 | 0.09 | 0.53 | 0.59 |
| | FR | 33 | 0.64 | 0.49 | 0.09 | | |
| | CG | 33 | 0.52 | 0.51 | 0.09 | | |
| 42 (show) | CR | 33 | 0.94 | 0.24 | 0.04 | 3.74 | 0.03 |
| | FR | 33 | 0.91 | 0.29 | 0.05 | | |
| | CG | 33 | 0.73 | 0.45 | 0.08 | | |
| 43 (performance) | CR | 33 | 0.76 | 0.44 | 0.08 | 0.05 | 0.95 |
| | FR | 33 | 0.76 | 0.44 | 0.08 | | |
| | CG | 33 | 0.73 | 0.45 | 0.08 | | |
| 44 (suddenly) | CR | 33 | 0.85 | 0.36 | 0.06 | 0.71 | 0.49 |
| | FR | 33 | 0.79 | 0.42 | 0.07 | | |
| | CG | 33 | 0.73 | 0.45 | 0.08 | | |
| 45 (traditional) | CR | 33 | 0.48 | 0.51 | 0.09 | 0.28 | 0.76 |
| | FR | 33 | 0.55 | 0.51 | 0.09 | | |
| | CG | 33 | 0.45 | 0.51 | 0.09 | | |
| 46 (find) | CR | 33 | 0.94 | 0.24 | 0.04 | 2.82 | 0.06 |
| | FR | 33 | 0.91 | 0.29 | 0.05 | | |
| | CG | 33 | 0.76 | 0.44 | 0.08 | | |
| 47 (carefully) | CR | 33 | 0.88 | 0.33 | 0.06 | 1.44 | 0.24 |
| | FR | 33 | 0.97 | 0.17 | 0.03 | | |
| | CG | 33 | 0.85 | 0.36 | 0.06 | | |
| 48 (hesitate) | CR | 33 | 0.85 | 0.36 | 0.06 | 0.40 | 0.67 |
| | FR | 33 | 0.91 | 0.29 | 0.05 | | |
| | CG | 33 | 0.91 | 0.29 | 0.05 | | |
| 49 (disturb) | CR | 33 | 0.85 | 0.36 | 0.06 | 1.78 | 0.17 |
| | FR | 33 | 0.97 | 0.17 | 0.03 | | |
| | CG | 33 | 0.94 | 0.24 | 0.04 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

As examined by the Fisher's LSD test, both experimental groups outperformed the Control group on Item 42, *show*, significantly (CR=0.94, FR=0.91, CG=0.73, CR vs. CG, $p=0.01$, FR vs. CG, $p=0.03$), and there was no significant difference between the two experimental groups ($p=0.72$) (see Table 4.24 below). On Item 46, *find*, only the Complementary Reading group significantly outperformed the Control group (CR=0.94, CG=0.76, $p=0.03$). On the same item, there is a practical difference between the Free Reading group and the Control group (FR=0.91, CG=0.76), although the difference is not statistically significant ($p=0.07$). There was no significant difference between the two experimental

Table 4. 24 The results of multiple comparisons by the Fisher's LSD test in the vocabulary section in the post-test

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|---------------------|-----------|-----------|-----------------------|------------|------|
| 41 (useless) | CR | FR | -0.03 | 0.12 | 0.81 |
| | | CG | 0.09 | 0.12 | 0.46 |
| | FR | CR | 0.03 | 0.12 | 0.81 |
| | | CG | 0.12 | 0.12 | 0.32 |
| | CG | CR | -0.09 | 0.12 | 0.46 |
| | | FR | -0.12 | 0.12 | 0.32 |
| 42 (show) | CR | FR | 0.03 | 0.08 | 0.72 |
| | | CG | 0.21 | 0.08 | 0.01 |
| | FR | CR | -0.03 | 0.08 | 0.72 |
| | | CG | 0.18 | 0.08 | 0.03 |
| | CG | CR | -0.21 | 0.08 | 0.01 |
| | | FR | -0.18 | 0.08 | 0.03 |
| 43 (performance) | CR | FR | 0.00 | 0.11 | 1.00 |
| | | CG | 0.03 | 0.11 | 0.78 |
| | FR | CR | 0.00 | 0.11 | 1.00 |
| | | CG | 0.03 | 0.11 | 0.78 |
| | CG | CR | -0.03 | 0.11 | 0.78 |
| | | FR | -0.03 | 0.11 | 0.78 |
| 44 (suddenly) | CR | FR | 0.06 | 0.10 | 0.55 |
| | | CG | 0.12 | 0.10 | 0.24 |
| | FR | CR | -0.06 | 0.10 | 0.55 |
| | | CG | 0.06 | 0.10 | 0.55 |
| | CG | CR | -0.12 | 0.10 | 0.24 |
| | | FR | -0.06 | 0.10 | 0.55 |
| 45 (traditional) | CR | FR | -0.06 | 0.12 | 0.63 |
| | | CG | 0.03 | 0.12 | 0.81 |
| | FR | CR | 0.06 | 0.12 | 0.63 |
| | | CG | 0.09 | 0.12 | 0.47 |
| | CG | CR | -0.03 | 0.12 | 0.81 |
| | | FR | -0.09 | 0.12 | 0.47 |
| 46 (find) | CR | FR | 0.03 | 0.08 | 0.71 |
| | | CG | 0.18 | 0.08 | 0.03 |
| | FR | CR | -0.03 | 0.08 | 0.71 |
| | | CG | 0.15 | 0.08 | 0.07 |
| | CG | CR | -0.18 | 0.08 | 0.03 |
| | | FR | -0.15 | 0.08 | 0.07 |
| 47 (carefully) | CR | FR | -0.09 | 0.07 | 0.22 |
| | | CG | 0.03 | 0.07 | 0.68 |
| | FR | CR | 0.09 | 0.07 | 0.22 |
| | | CG | 0.12 | 0.07 | 0.11 |
| | CG | CR | -0.03 | 0.07 | 0.68 |
| | | FR | -0.12 | 0.07 | 0.11 |
| 48 (hesitate) | CR | FR | -0.06 | 0.08 | 0.44 |
| | | CG | -0.06 | 0.08 | 0.44 |
| | FR | CR | 0.06 | 0.08 | 0.44 |
| | | CG | 0.00 | 0.08 | 1.00 |
| | CG | CR | 0.06 | 0.08 | 0.44 |
| | | FR | 0.00 | 0.08 | 1.00 |
| 49 (disturb) | CR | FR | -0.12 | 0.07 | 0.07 |
| | | CG | -0.09 | 0.07 | 0.18 |
| | FR | CR | 0.12 | 0.07 | 0.07 |
| | | CG | 0.03 | 0.07 | 0.65 |
| | CG | CR | 0.09 | 0.07 | 0.18 |
| | | FR | -0.03 | 0.07 | 0.65 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

groups on Item 46 (CR=0.94, FR=0.91, $p=0.71$). The two experimental groups' better performance on the highest-frequency words in the post-test was consistent with the research literature that how well the readers acquire a word is proportional to its frequency.

The second interesting finding emerges from Item 48 *hesitate* and Item 49 *disturb*. On these two items, the Free Reading group and the Control group outperform the Complementary Reading group in terms of mean score although the difference was not statistically significant, and the Free Reading group and the Control group performed equally well on this item by obtaining the same mean score. A practical explanation for this finding is that the participants in the Free Reading group and the Control group had intensive grammar and translation exercises on these two words, because they are in the vocabulary list attached to the textbook and understandably explicitly taught by the teachers. The participants in the Complementary Reading group, who did not do grammar and translation exercises but extensive reading after school, were reasonably exposed to these two words less.

The three groups' performance on these two items was consistent with Laufer's three studies noted earlier in this section. The findings from the three studies showed that engaging with the words more "consciously" produced better gains in vocabulary learning than through reading. The value of findings from the present study as well as Laufer's experiments is to show the significance of learners acquiring words through explicit instruction or "conscious" tasks. The limitation of acquiring words through explicit instruction or "conscious" tasks is that the huge number of words that learners are supposed to know make impossible the acquisition or learning of every new word by means of explicit instruction or "conscious" task. Most L1 direct vocabulary instruction in classroom does not involve more than 300 to 500 words per academic year (Stahl & Nagy, 2006), and in an L2 setting a fairly intensive instruction of vocabulary (12-20 hours per week) can teach 2,000 words directly to the learners (Grabe, 2009). Given that the general consensus on how many words there are in English is that over 100,000 words are in active use (Nagy, 1997; Nation, 2001; Schmitt, 2000), it is certainly beyond the scope of explicit instruction or "conscious" task

to teach the 100,000 to the learners, and that is why Nation (2001) points out that balanced vocabulary instruction should comprise input, instruction, output and fluency development, as noted earlier in this section.

The overview of the vocabulary instruction practice in Chinese secondary school made earlier in this section has shown that the second and the third strands in Nation's framework, the language-focused or form-focused instruction and meaning-focused output, are well incorporated into in-class vocabulary instruction, while the other two fail to be clearly reflected. It is extensive reading that plays a crucial and irreplaceable role in learning from comprehensible meaning-focused input and fluency development.

Findings from *hesitate* and *disturb* also reveal a challenge to the testing and assessment of vocabulary gains through extensive reading. Obviously the locally developed tests are heavily teaching-oriented. For the complementary reading participants who spent more time reading but less time on explicit instruction of isolated lexical items, they were at a disadvantage in the test.

To summarize, the findings in the section show that explicit instruction and extensive reading are complementary to each other in terms of vocabulary learning in that they represent two major processes by which learners add new words and expand vocabulary knowledge. The Grammar-Translation Method involves systematic procedures to memorize the forms and meanings of words, usually based on lists of English words with their Chinese equivalents or translated meanings, which is of particular value for the learning or acquisition of less frequent words. Extensive reading promotes individual vocabulary learning, by which readers acquire new words or expand knowledge of known word incrementally as the readers encounter the words in context in their extensive reading, and it also facilitate the fluency development of known words.

4.4.2 Extensive reading and grammar acquisition

Prior to interpreting and discussing the test results in the grammar section, it should be noted here that grammar is defined as a language knowledge component rather than a skill in the present study. There are three reasons for

defining grammar as a language knowledge component in the present study. The first reason is that the teaching of grammar is described as instruction of rules or language knowledge in the consecutive national syllabuses for English language teaching as shown in Table 4.25 below. Further, the grammar sections in the pre-, mid- and post-tests without exception employed a multiple-choice format. The students were requested to complete a sentence by choosing from the four choices provided to complete a sentence. The tests did not provide contexts but only single sentences. In essence, the students were not involved in communicative tasks because they did not know who the intended audience was. They have to focus on the structure rather than meaning in order to locate the correct answer from the four options. The third reason for defining grammar as a language knowledge component rather than a skill is that grammar was not tested as part of a test of writing or speaking, which means the accuracy and fluency of grammar was not tested.

Table 4. 25 Guidelines for teaching grammar in the national syllabuses for English language teaching

| Syllabus | Guidelines for teaching grammar |
|----------|---|
| 1978 | The students are expected to have a command of basic phonetics and grammar rules. |
| 1986 | Give detailed instruction of basic language knowledge. |
| 1988 | Strengthen basic language knowledge and training. |
| 1990 | The same as the 1988 syllabus. |
| 1993 | Consolidate and expand students' basic language knowledge. |
| 2000 | The same as the 1993 syllabus. |

Returning from the nature of the grammar test and the definition of grammar in the present study, grammar is recognized as being fundamental to language. Definitions of the word “grammar” are manifold: they can range from a “theory of language” to “a description of the syntactic structure of a particular language” (Pachler & Field, 1997, p.145); it can also be thought of as “a type of rule-governed behavior” (Celce- Murica & Hilles, 1988, p.16).

Of the many issues surrounding grammar in a second or foreign language classroom, perhaps the most controversial is whether to teach it or not. From an historical perspective, this controversy comes as somewhat of a surprise: for thousands of years, grammar was the center of language pedagogy (Rutherford,

1987). Language teaching was equated with grammar teaching and grammar was used as content as well as an organizing principle for developing curriculum and language teaching materials (Celce-Murcia, 2001). The nature of this controversy revolves around the question of whether or not the instruction of grammar rules facilitates learners' improvement in the target language proficiency.

Recent research in second language acquisition has led to a radical rethinking of the importance of grammar. Many researchers now believe that grammar teaching should *not* be ignored in second language classrooms. Language teaching professionals have also become increasingly aware that the importance of grammar instruction in language teaching and learning should not be underestimated.

The second controversial issue surrounding the teaching of grammar is what to teach first. The Natural Order Hypothesis proposed by Krashen (1981) suggests that learners, regardless of their language background, acquire grammatical constructions in a certain language in a predicable order. Some grammar structures in a certain language are more likely to be acquired earlier, and other structures in this language later. However, this predicable order is not reflected in the teaching of grammar. General educational principles suggest we start with the simple and move toward the more difficult, but structures that seem uncomplicated may in fact be quite difficult for some students to internalize. For example, most adult learners still omit the obligatory –s from third person singular verbs in the present tense long after they have mastered “more complicated” verb forms. Obviously, “simple” and “complex” are meaningless terms if they are not defined from the learners' point of view. Although there have been some attempts to identify a “difficulty” or “acquisition” order in English, the resulting sequence usually deals only with elementary structures and cannot possibly account for every structure in the language.

The third controversy on the teaching of grammar is the debate on an explicit versus an implicit view of grammar instruction. According to Kelly (1969), this controversy has existed since the beginning of language teaching. So far no consensus has been reached on this debate. Some theorists advocate formal and

systematic attention to isolate linguistic features, whereas others promote whole language experience in the target language, which bears strong resemblance to how children acquire their first language (Long & Richards, 1998). On this debated, the teachers, who are directly engaged in classroom teaching, hold two opposing viewpoints, supporting either exposure to the target language or explicit instruction as the major source of grammar competence (Garrett, 1986).

Although no agreement has been reached on them, the three controversies noted above did not lead to dispute in mainland China. As shown in Table 4.25, from 1978, the importance of grammar was highlighted in six consecutive national syllabuses for English language teaching, and since then the Grammar-Translation Method dominates English language teaching in classroom.

Under the guidelines of the national English language teaching syllabuses, grammar has been afforded a very high profile in the highest-stakes NMET. As has been shown in Table 1.4 in Chapter 1, grammar has been an integral part of the NMET since 1978. Due to the decentralization policy of the State Education Committee that every province, autonomous region and municipality be encouraged to develop its own version of NMET to cater for the local education situation, a uniform version of the NMET was no longer developed after 1990. In the localized NMETs, grammar is still a crucial component until today.

The direct washback effect of the syllabuses and the NMET on English language teaching in the classroom is that the Grammar-Translation Method dominates English language teaching in classroom. As its name suggests, the Grammar-Translation Method lays great emphasis on grammar and translation, and the essence is to build accuracy and in turn to cultivate learners' ability to construct correct sentences (Griffiths & Parr, 2001).

A typical lesson using the Grammar-Translation Method in mainland China was observed by Zuo (2002):

1. Review: Dictation of words; individual and collection recitation of a paragraph in the text.

2. Teaching of new words: The teacher explained the new words in the text verbatim. The students read the new words after the teacher. Afterwards, the teacher proposed a Chinese word, and the students answer the English equivalent of the Chinese word proposed by the teacher.
3. Teaching of grammar rules (simple progressive tense): After explaining the rules of the simple progressive tense, the teacher wrote the standard forms of the verbs in the text on the blackboard, and requested the students to change the standard forms into present participles and then simple progressive tense.
4. Teaching of the text: The teacher read the text sentence by sentence, and requested the students to analyze the grammar structure of the sentences and translate them into Chinese. The teacher corrected the students' mistakes at any time.
5. Consolidation of the text: The students read the text after the teacher sentence by sentence. Then the teacher raised questions about the text, and the students answered the questions based on the content of the text.
6. Assignment: Recitation of new words; grammar exercise: simple progressive tense; recitation of a paragraph in the text.

(Zuo, 2002, p.39)

The teaching practice of the Grammar-Translation Method bears great resemblance to the popular Presentation-Practice-Produce (P-P-P) teaching style. Byrne (1986) explores typical steps of P-P-P. First, the teacher *presents* new language items. Second, learners *practice* the items through drills, and then *produce* the language for themselves. From the 1980s onwards, P-P-P has been recommended as a useful teaching procedure (Harmer, 2007), as it has a logic that is appealing to teachers and a learner in that it reflects a notion of practice makes perfect (Carless, 2009). Swan (2005) also recommends PPP as a useful routine for presenting and practicing structural features under semi-controlled conditions. It also allows the teacher to control the content and pace of the lessons (Thornbury, 1999).

The emphasis on grammar shows that English language teaching in Chinese secondary school is based on the assumption that “the major problem in learning a second or foreign language is learning its structure and that this aspect of

language must receive exclusive attention” (Nassaji & Fotos, 2011, p. 2). If so, it is even more necessary to discuss whether the Grammar-Translation Method is ideal for teaching and learning grammar.

One of the major assumptions underlying traditional grammar-based approaches is that language consists of a series of grammatical forms and structures that can be acquired successively (Nassaji & Fotos, 2011). Grammar teaching is viewed as a deductive and linear presentation of these rules. It is believed that through such presentations of grammar forms, learners are able to develop the kind of knowledge they need for spontaneous language use, but practice of these rules is excluded by the strictly controlled classroom teaching because the Grammar-Translation Method teaches language through rules rather than use.

In addition to the “rules” and “use” contradiction, another conflict in the Grammar-Translation Method is “acquisition” and “learning”. This method allows for little opportunity for acquisition but solely relies on learning. In the past several years many ESL professionals have come to assume that second language learning bears strong resemblance first language acquisition. Their position on the “acquisition” and “learning” issue is that acquisition is more important than learning (e.g., Krashen, 1993).

Once it is agreed that acquisition is more important than learning, questions concerning acquisition rise immediately. For acquisition to occur, the prerequisite is that comprehensible input through which the learners acquire is made available, and preferably, in a low affective environment. As discussed in Chapter 1 and 2, extensive reading has been investigated as an optimal source of comprehensible input, particularly in an EFL context where English is not the language of daily communication. There have been numerous studies on the significance of extensive reading in a comprehensive range of areas, from language knowledge, receptive skills, productive skills and even vernacular and other school subjects (see Section 2.1 for review of benefits of extensive reading from empirical studies and Section 2.5 for benefits of extensive reading from large-scale extensive reading programs). As noted from the existing body of research, the positive role that extensive reading plays in grammar acquisition is

well-recognized among researchers. Findings from the extensive reading programs worldwide also support the positive relationship between extensive reading and acquisition of grammar competence.

In REAP, which has been introduced in detail in Section 2.5, the participants were tested on a wide range of English language tests on word recognition, oral language, writing, reading, vocabulary, grammar, etc (Mok, 1994). The grammar test employed a multiple-choice format that assesses the participants on syntactic structure. The test result showed that the REAP participants were significantly better than the non-REAP participants on grammar (REAP=11.0, non-REAP=8.9, $t=6.6$, $p<0.01$). In the Book Flood Project in Fiji, which has also been introduced in detail in Section 2.5, the two experimental groups, which experienced two different experimental treatments, namely the Silent Reading Treatment and the Shared Reading Treatment, outperformed the control group significantly in the open-ended English structures test ($p<0.01$) (Elley & Mangubhai, 1983, Elley, 1991).

These two extensive reading programs lend the following two implications for the present study:

1. Extensive reading does have a role to play in grammar acquisition.
2. Differences in extensive reading treatments did not seem to significantly affect the effect of extensive reading on grammar acquisition.

According to the research design of the present study, the three groups experienced three models of grammar instruction and acquisition. The participants in the Control group were explicitly taught the grammar rules in class by their English teacher. They had to finish their grammar and translation exercises after school, but they did not read extensively. In other words, they were taught using the Grammar-Translation Method, and were not given additional comprehensible input other than their textbooks. As for the Complementary Reading group, participants in this group were explicitly taught the grammar rules in class by the researcher. This was similar to the Control group, because grammar was a crucial part for the Complementary Reading group in teaching content as required in the teaching syllabus, and absence of

explicit instruction of grammar was strictly forbidden. However, they did not do grammar or translation exercises after school; instead, extensive reading was their daily assignment, and they read about 5,000 words every day after school. Participants in the Free Reading group were explicitly taught the grammar rules in class by their class English teachers as well. They had to finish their grammar and translation exercises after school, and read out of class when they had time. There was no specific amount of reading required.

As shown in Table 4.26, the pre-test showed no statistically significant difference in grammar among the three groups ($F=0.96$, $p=0.39$). It showed that the three groups joined this extensive reading program with approximately equal grammar competence. The situation changed in the mid-test. Judging from the mean score, the Free Reading group performed the best among the three on grammar in the mid-test (CR=14.15, FR=15.15, CG=13.97), but there is no significant main effect for the treatments as a factor in the grammar section in the mid-test ($F=2.16$, $p=0.12$). In the post-test, the Free Reading group still had the advantage in grammar (CR=15.88, FR=16.64, CG=15.03) judging from the mean score, there is no significant main effect for the treatments as a factor in grammar in the post-test either ($F=2.68$, $p=0.07$).

Table 4. 26 Descriptive statistics and results of ANOVA from the grammar sections in the pre-, mid- and post-tests

| Tests | Group | N | Mean | SD | F | Sig. |
|-----------|-------|----|-------|------|------|------|
| Pre-test | CR | 33 | 3.61 | 1.20 | 0.96 | 0.39 |
| | FR | 33 | 3.97 | 0.95 | | |
| | CG | 33 | 3.73 | 1.10 | | |
| Mid-test | CR | 33 | 14.15 | 2.62 | 2.16 | 0.12 |
| | FR | 33 | 15.15 | 2.55 | | |
| | CG | 33 | 13.97 | 2.27 | | |
| Post-test | CR | 33 | 15.88 | 2.67 | 2.68 | 0.07 |
| | FR | 33 | 16.64 | 2.22 | | |
| | CG | 33 | 15.03 | 3.44 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Table 4.27 below illustrated the multiple comparisons by Fisher's LSD test between the three groups in the grammar sections in the pre-, mid- and post-tests. It is worth noting that there is a practical difference between the Free Reading group and the Control group in the mid-test although the difference is not statistically significant ($p=0.06$), and in the post-test the Fisher's LSD test

detected a statistically significant difference between the Free Reading group and the Control group ($p=0.02$). It should be stressed that the difference between the Free Reading group and the Control group in the post-test was the only statistical difference in the grammar sections in the post-tests. There was no statistical difference between the two experimental groups ($p=0.28$) or between the Complementary Reading group and the Control group ($p=0.22$) in the grammar section in the post-test.

Table 4. 27 The results of multiple comparisons by the Fisher's LSD test in the grammar sections in the pre-, mid- and post-tests

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| Pre-test | CR | FR | -0.36 | 0.28 | 0.20 |
| | | CG | -0.30 | 0.28 | 0.29 |
| | FR | CR | 0.36 | 0.28 | 0.20 |
| | | CG | 0.06 | 0.28 | 0.83 |
| | CG | CR | 0.30 | 0.28 | 0.29 |
| | | FR | -0.06 | 0.28 | 0.83 |
| Mid-test | CR | FR | -1.00 | 0.61 | 0.11 |
| | | CG | 0.18 | 0.61 | 0.77 |
| | FR | CR | 1.00 | 0.61 | 0.11 |
| | | CG | 1.18 | 0.61 | 0.06 |
| | CG | CR | -0.18 | 0.61 | 0.77 |
| | | FR | -1.18 | 0.61 | 0.06 |
| Post-test | CR | FR | -0.76 | 0.69 | 0.28 |
| | | CG | 0.85 | 0.69 | 0.22 |
| | FR | CR | 0.76 | 0.69 | 0.28 |
| | | CG | 1.61 | 0.69 | 0.02 |
| | CG | CR | -0.85 | 0.69 | 0.22 |
| | | FR | -1.61 | 0.69 | 0.02 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Discussion 1: What is the significance of acquiring grammar through extensive reading?

What the three groups shared in common in their treatments was the explicit instruction of the grammar rules, and thus the statistical difference in the post-test can only be attributed to extensive reading. In the post-test, the statistically significant difference appeared between the Free Reading group and the Control group. This was evidence in favor of extensive reading, since the only extraneous variable between the two groups was extensive reading on condition that other extraneous variables had been strictly controlled. Another finding that may bring out the limitation of the traditional Grammar-Translation Method is that the Complementary Reading group did not do any grammar and translation exercises but it did seem to be slightly better at grammar than the Control group although the difference was not statistically significant (CR=15.88, CG=15.03, $p=0.22$).

The traditional Grammar-Translation Method has long dominated the English language classroom. Both the teachers and students stuck to investing a huge amount of time discussing grammar in class and doing grammar and translation exercise after class. However, the findings did not support the effectiveness of the Grammar-Translation Method, i.e. the explicit instruction of grammar rules in class followed by the grammar and translation exercises after school. The significance of acquiring grammar through comprehensible input is at least as effective as the Grammar-Translation Method with the additional merit of being more interesting to the learners.

In addition, the empirical research supporting the effect of extensive reading on grammar has provided evidence that the students' grammar did improve through extensive reading. For example, in a 90-hour extensive reading program based on graded readers, twenty-five ESL learners in Pakistan made significant gains in terms of grammar fluency and accuracy after extensive reading (Hafiz & Tudor, 1989). The weakness of extensive reading on grammar competence, if there is any, is that the participants in this study did not improve in range of structures used.

From the analysis and empirical studies discussed above, a conclusion can be reached that there is a positive relationship between extensive reading and grammar acquisition. On the other hand, it should be noted that the evidence produced by the empirical studies was not enough to explain the whole picture of how extensive reading positively influences grammar. Hafiz and Tudor's (1989) study did not cover a comprehensive range of grammatical structures. In the present study, although the Free Reading group did significantly better in grammar, it is notable that the result came from only twenty multiple choices. The twenty choices were far from representative of the whole grammar and structure rules.

Discussion 2: How did extensive reading affect grammar acquisition?

In response to this argument that the twenty choices were far from representative of the whole grammar and structure rules, it is certainly necessary to revisit the three tests. The ten grammar items in the pre-test included the following: modal

verb, pronoun, adjective, adverb, adverbial clause, compound sentence, verb, and attributive clause (the last two items tested twice) (see Appendix A). The grammar section in the mid-test tested conjunction, compound sentence, non-finite verb, modal verb, object clause, adverbial clause, attributive clause, present perfect tense and past continuous tense (see Appendix B). The grammar items in the post-test included the following: preposition, pronoun, gerund as subject, gerund as object, infinitive as adverbial, participle as attributive, past continuous tense, simple past tense, attributive clause, adverbial clause, object clause, subject clause, exclamatory sentence, passive voice, subject-verb agreement and inversion (see Appendix C).

The pre-test mainly covered the grammar knowledge the students learned in junior middle school. As required in the national junior middle school English syllabus, the junior middle school graduates were required to have basic English language knowledge. As illustrated in Table 4.28 below, the grammar syllabus in the Shanghai English Curriculum Standards for Primary and Junior Middle School (2004) lists the grammar knowledge that students are required to learn in detail: parts of speech and sentence knowledge in terms of components of sentences, types of sentences, simple sentences, complex sentences and compound sentences. The grammar rules tested in the pre-test mainly covered parts of speech including verb, pronoun, adjective and adverbial, complex sentences including attributive clause and adverbial clause, and imperative and compound sentences. The pre-test was designed to strictly adhere to the syllabus, aiming to test the test takers' grammar knowledge learned in junior middle school.

Another trend that can be noted from Table 4.28 is that the content of the three tests shared very little in common. The one reason for this was that no more than twenty items were devoted to grammar in each test, so it would be impossible to cover every grammar rule in a single test. As discussed in the previous paragraph, the content of the grammar section was highly consistent with the syllabus and the classroom teaching. In this sense, the three tests were not typical representatives of the English grammar system, but typical representatives of the grammar syllabus and the contemporary explicit instruction of grammar in

Table 4. 28 The grammar syllabus in the Shanghai English Curriculum Standards for Primary and Junior Middle School (2004)

| Description | | | OBW | Pre | Mid | Post |
|-------------------|-----------------|-----------|--|--------------------------|-----|------|
| Sentence segments | Parts of speech | noun | countable/uncountable nouns | • | | |
| | | | plural forms of nouns | • | | |
| | | | proper nouns | • | | |
| | | | possessives | • | | |
| | | adjective | adjectives used as attributive, predicative and object complement | • | • | |
| | | | formation of comparative/superlative forms: er, est; more, the most. | • | | |
| | | | sentence patterns: | | | |
| | | | ● as + original form of adjective + as.., | • | | |
| | | | ● not as (so) + original form of adjective+ as... | | | |
| | | | ● comparative form of adjective + than, | | | |
| | | | ● the + superlative form of adjective + ... in (of) | | | |
| | | adverb | adverbs used to indicate time, place, manner and degree, etc. | • | • | • |
| | | | doubt adverbs: when, where, how. | • | | |
| | | | formation of comparative/superlative forms: er, est; more, the most. | • | | |
| | | | sentence patterns: | | | |
| | | | ● as + original form of adjective + as.., | • | | |
| | | | ● not as (so) + original form of adjective+ as... | | | |
| | | | ● comparative form of adjective + than, | | | |
| | | | ● the + superlative form of adjective + ... in (of) | | | |
| | | verb | transitive verb | • | | |
| | | | intransitive verb | link verbs | • | |
| | | | | auxiliary verbs | • | |
| | | | | modal verbs | • | • |
| | | | tense | simple present tense | • | |
| | | | | simple past tense | • | • |
| | | | | simple future tense | • | |
| | | | | present continuous tense | • | |
| | | | | present perfect tense | • | • |
| | | | | past continuous tense | • | • |
| | | | | past perfect tense | • | • |
| | | | | past future tense | | |

| | | | | | | | | |
|----------|-------------------------------|--------------------------|---|---|---|---|---|---|
| | | | voice | active voice | • | | | |
| | | | | passive voice (in simple present and past tense/in simple present tense with modal verbs) | • | | | |
| | | | infinite | infinite used as subject | • | | | |
| | | | | infinite used as object | • | | • | |
| | | | | infinite used as object complement | • | | | |
| | | | | infinite used as attributive | • | | | |
| | | | | infinite used as adverbial | • | | | • |
| | | | | infinite used after when/ where/ how/ what/which | • | | | |
| | | | Pronoun | personal pronouns in nominative and accusative forms | • | | | • |
| | | | | possessive pronouns in adjective and noun forms | • | | | |
| | | | | reflexives: myself, himself, ourselves, etc. | • | | | |
| | | | | demonstratives: this, that, these, those. | • | | | |
| | | | | indefinite pronouns: some, any, no, etc. | • | • | | |
| | | | | interrogative pronouns: what, who, whose, which, etc. | • | | | |
| | | | article | definite article | • | | | |
| | | | | indefinite article | • | | | |
| | | | | zero article | • | | | |
| | | | numeral | cardinal numbers | • | | | |
| | | | | ordinals numbers | • | | | |
| | | | Preposition | above/among/around/during/except inside/outside/without | • | | | • |
| | | | conjunction | and/or | • | | • | |
| | | | interjection | / | • | | | |
| | Word formations | combination | e.g. classroom. Something, reading-room | • | | | | |
| | | deviation | e.g. worker, drawing, quickly | • | | | | |
| | | transformation | e.g. record (n.)—record (v.) | • | | | | |
| Sentence | Components of sentences | subject | | • | | | | |
| | | verb | | • | | | | |
| | | predicative | | • | | | | |
| | | object (direct/indirect) | | • | | | | |

| | | | | | | |
|--|-----------------------------|---|---|---|---|---|
| | | object complement | • | | | |
| | | attribute | • | | | |
| | | adverbial | • | | | |
| | Types of sentences | declarative sentence (positive/negative) | • | | | |
| | | interrogative sentence (general/special/alternative/tag question) | • | | | |
| | | imperative sentence (positive/negative) | • | • | | |
| | | exclamatory sentence | • | | | • |
| | Simple sentences | subject+verb (SV) | • | | | |
| | | subject+verb+predicative (SVP) | • | | | |
| | | subject+verb+object (SVO) | • | | | |
| | | subject+verb+ indirect object+ direct object (SVOiOd) | • | | | |
| | | subject+verb+object+complement (SVOC) | • | | | |
| | Complex sentences | the object clause | • | | • | • |
| | | the adverbial clause | • | • | • | • |
| | | the attributive clause | • | • | • | • |
| | Compound sentence (and/but) | | • | | • | |

(adapted from the Shanghai English Curriculum Standards for Primary and Junior Middle School, 2004, p. 64-68)

classroom. This was the common practice of grammar teaching and testing in mainland China.

Thus how extensive reading affected grammar acquisition can be investigated in two complementary ways: (1) how the two experimental groups performed on the grammar items which had been explicitly taught in class, and (2) how they performed on grammar items which had not been explicitly taught in class during the program.

Whether the grammar items had been explicitly taught was judged according to the actual situation of grammar teaching during the academic year. After being admitted into senior high school, the students are to be explicitly taught three major grammars in their first year of senior high school: (1) non-finite verbs, including gerund, infinitive, participle and their specific usages such as gerund as subject/object, (2) complex sentences including noun clauses, the attributive clause and the adverbial clause, and (3) nearly all tenses in English. Other grammatical constructions included modal verbs and inversion. Briefly speaking, the students will have been explicitly taught all basic English grammar rules by the end of the first academic year in senior high school. The contrast between the grammar sections in the mid- and post-tests reveals that the grammar section in the post-test was more comprehensive than that in the mid-test. The explanation for this difference is that the mid-test, i.e. the final term test for the first academic term, usually serves as a monitoring test at the end of the first academic term, while the post-test serves as an evaluation test at the end of the first academic year. The teaching progress of individual schools in the same school district may vary by the end of the first academic term, but they must have finished teaching required grammar rules by the end of the first academic year. That was why the grammar section in the post-test was more comprehensive and evaluative.

According to the description of grammar teaching in the whole academic year, the twenty grammar items in the post-test can be grouped into two categories for our discussion purpose: grammatical constructions that have not been explicitly taught, i.e., Items 21, 22, 24 and 26 and grammatical constructions that have been explicitly taught, i.e. sixteen items excluding Items 21, 22, 24 and 26.

Table 4.29 presents the descriptive statistics and results of ANOVA from the individual items in the grammar section in the post-test. As shown in Table 4.29, there is a significant main effect for the treatments as a factor on Item 21 ($F=3.01$, $p=0.05$) and Item 22 ($F=3.35$, $p=0.04$). A practical difference between the three groups is also found on Item 26, although the difference is not statistically significant ($F=2.86$, $p=0.06$). There is no significant main effect for the treatments as a factor on Item 24 ($F=1.13$, $p=0.33$).

In Table 4.30 below, the differences between each group was further examined. On Items 21 and 22, there was a statistically significant difference between the Free Reading group and the Control Group (Item 21: FR=1.00, CG=0.85, $p=0.02$; Item 22: FR=0.73, CG=0.42, $p=0.01$). There was no statistically significant difference between the two experimental groups on these two items (Item 21: CR=0.94, FR=1.00, $p=0.33$; Item 22: CR=0.52, FR=0.73, $p=0.08$). There was no statistically significant difference between the three groups on Item 24.

The situation of Item 26 was different from that on Item 21, 22 and 24. On Item 21, 22 and 24, without exception, the two experimental groups outperformed the control group. Also on these three items, without exception as well, the Free Reading group outperformed the Complementary Reading group. On Item 26, although the Free Reading group outperformed both the Complementary Reading group and the Control group (CR=0.85, FR=1.00, CG=0.85, CR vs. FR, $p=0.04$; CG vs. FR, $p=0.04$; CR vs. CG, $p=1.00$), the Complementary Reading group did not, however, outperform the Control group.

This exception for Item 26 prevents us from making the blanket statement that extensive reading is effective in terms of incidental grammar acquisition. On one hand, the Free Reading group outperformed the Control group significantly on the grammar rules that were not explicitly taught; on the other hand, there was no statistically significant difference between the Complementary Reading group and the Control group on the grammar rules that were not explicitly taught.

The situation in the other group of items, namely the grammar rules which were explicitly taught, is more complex. On these sixteen items, no group obtained

Table 4. 29 Descriptive statistics and results of ANOVA from the grammar section in the post-test

| Item | Group | N | Mean | SD | Std. Error | F | Sig. |
|--|-------|----|------|------|------------|------|------|
| 21 preposition | CR | 33 | 0.94 | 0.24 | 0.04 | 3.01 | 0.05 |
| | FR | 33 | 1.00 | 0.00 | 0.00 | | |
| | CG | 33 | 0.85 | 0.36 | 0.06 | | |
| 22 pronoun | CR | 33 | 0.52 | 0.51 | 0.09 | 3.35 | 0.04 |
| | FR | 33 | 0.73 | 0.45 | 0.08 | | |
| | CG | 33 | 0.42 | 0.50 | 0.09 | | |
| 23 attributive clause | CR | 33 | 0.85 | 0.36 | 0.06 | 0.28 | 0.76 |
| | FR | 33 | 0.85 | 0.36 | 0.06 | | |
| | CG | 33 | 0.79 | 0.42 | 0.07 | | |
| 24 exclamatory sentence | CR | 33 | 0.61 | 0.50 | 0.09 | 1.13 | 0.33 |
| | FR | 33 | 0.70 | 0.47 | 0.08 | | |
| | CG | 33 | 0.52 | 0.51 | 0.09 | | |
| 25 gerund | CR | 33 | 0.91 | 0.29 | 0.05 | 1.66 | 0.20 |
| | FR | 33 | 0.88 | 0.33 | 0.06 | | |
| | CG | 33 | 0.76 | 0.44 | 0.08 | | |
| 26 conjunction | CR | 33 | 0.85 | 0.36 | 0.06 | 2.86 | 0.06 |
| | FR | 33 | 1.00 | 0.00 | 0.00 | | |
| | CG | 33 | 0.85 | 0.36 | 0.06 | | |
| 27 tense (past continuous tense) | CR | 33 | 0.67 | 0.48 | 0.08 | 0.19 | 0.83 |
| | FR | 33 | 0.73 | 0.45 | 0.08 | | |
| | CG | 33 | 0.73 | 0.45 | 0.08 | | |
| 28 non-finite verb | CR | 33 | 0.76 | 0.44 | 0.08 | 0.79 | 0.46 |
| | FR | 33 | 0.64 | 0.49 | 0.09 | | |
| | CG | 33 | 0.76 | 0.44 | 0.08 | | |
| 29 attributive clause | CR | 33 | 0.79 | 0.42 | 0.07 | 1.17 | 0.32 |
| | FR | 33 | 0.64 | 0.49 | 0.09 | | |
| | CG | 33 | 0.64 | 0.49 | 0.09 | | |
| 30 complex sentence (object clause) | CR | 33 | 0.97 | 0.17 | 0.03 | 1.00 | 0.37 |
| | FR | 33 | 1.00 | 0.00 | 0.00 | | |
| | CG | 33 | 1.00 | 0.00 | 0.00 | | |
| 31 complex sentence (subject clause) | CR | 33 | 0.91 | 0.29 | 0.05 | 3.20 | 0.05 |
| | FR | 33 | 1.00 | 0.00 | 0.00 | | |
| | CG | 33 | 1.00 | 0.00 | 0.00 | | |
| 32 tense (simple past tense) | CR | 33 | 0.73 | 0.45 | 0.08 | 1.36 | 0.26 |
| | FR | 33 | 0.79 | 0.42 | 0.07 | | |
| | CG | 33 | 0.61 | 0.50 | 0.09 | | |
| 33 attributive clause | CR | 33 | 0.85 | 0.36 | 0.06 | 0.84 | 0.43 |
| | FR | 33 | 0.94 | 0.24 | 0.04 | | |
| | CG | 33 | 0.85 | 0.36 | 0.06 | | |
| 34 participle as attributive | CR | 33 | 0.76 | 0.44 | 0.08 | 0.65 | 0.53 |
| | FR | 33 | 0.82 | 0.39 | 0.07 | | |
| | CG | 33 | 0.70 | 0.47 | 0.08 | | |
| 35 adverbial clause | CR | 33 | 0.82 | 0.39 | 0.07 | 0.06 | 0.94 |
| | FR | 33 | 0.82 | 0.39 | 0.07 | | |
| | CG | 33 | 0.79 | 0.42 | 0.07 | | |
| 36 subject-verb agreement | CR | 33 | 0.79 | 0.42 | 0.07 | 0.20 | 0.82 |
| | FR | 33 | 0.85 | 0.36 | 0.06 | | |
| | CG | 33 | 0.82 | 0.39 | 0.07 | | |
| 37 gerund as subject | CR | 33 | 0.88 | 0.33 | 0.06 | 2.35 | 0.10 |
| | FR | 33 | 0.91 | 0.29 | 0.05 | | |
| | CG | 33 | 0.73 | 0.45 | 0.08 | | |
| 38 attributive clause | CR | 33 | 0.88 | 0.33 | 0.06 | 1.44 | 0.24 |
| | FR | 33 | 0.97 | 0.17 | 0.03 | | |
| | CG | 33 | 0.85 | 0.36 | 0.06 | | |
| 39 attributive clause | CR | 33 | 0.64 | 0.49 | 0.09 | 0.29 | 0.75 |
| | FR | 33 | 0.61 | 0.50 | 0.09 | | |
| | CG | 33 | 0.55 | 0.51 | 0.09 | | |
| 40 inversion | CR | 33 | 0.79 | 0.42 | 0.07 | 0.84 | 0.44 |
| | FR | 33 | 0.88 | 0.33 | 0.06 | | |
| | CG | 33 | 0.76 | 0.44 | 0.08 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Table 4. 30 The results of multiple comparisons by the Fisher's LSD test in the grammar section in the post-test

| Item | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|---|-----------|-----------|-----------------------|------------|------|
| 21 preposition | CR | FR | -0.06 | 0.06 | 0.33 |
| | | CG | 0.09 | 0.06 | 0.15 |
| | FR | CR | 0.06 | 0.06 | 0.33 |
| | | CG | 0.15 | 0.06 | 0.02 |
| | CG | CR | -0.09 | 0.06 | 0.15 |
| | | FR | -0.15 | 0.06 | 0.02 |
| 22 pronoun | CR | FR | -0.21 | 0.12 | 0.08 |
| | | CG | 0.09 | 0.12 | 0.45 |
| | FR | CR | 0.21 | 0.12 | 0.08 |
| | | CG | 0.30 | 0.12 | 0.01 |
| | CG | CR | -0.09 | 0.12 | 0.45 |
| | | FR | -0.30 | 0.12 | 0.01 |
| 23 attributive clause | CR | FR | 0.00 | 0.09 | 1.00 |
| | | CG | 0.06 | 0.09 | 0.52 |
| | FR | CR | 0.00 | 0.09 | 1.00 |
| | | CG | 0.06 | 0.09 | 0.52 |
| | CG | CR | -0.06 | 0.09 | 0.52 |
| | | FR | -0.06 | 0.09 | 0.52 |
| 24 exclamatory sentence | CR | FR | -0.09 | 0.12 | 0.45 |
| | | CG | 0.09 | 0.12 | 0.45 |
| | FR | CR | 0.09 | 0.12 | 0.45 |
| | | CG | 0.18 | 0.12 | 0.14 |
| | CG | CR | -0.09 | 0.12 | 0.45 |
| | | FR | -0.18 | 0.12 | 0.14 |
| 25 gerund | CR | FR | 0.03 | 0.09 | 0.73 |
| | | CG | 0.15 | 0.09 | 0.09 |
| | FR | CR | -0.03 | 0.09 | 0.73 |
| | | CG | 0.12 | 0.09 | 0.17 |
| | CG | CR | -0.15 | 0.09 | 0.09 |
| | | FR | -0.12 | 0.09 | 0.17 |
| 26 conjunction | CR | FR | -0.15 | 0.07 | 0.04 |
| | | CG | 0.00 | 0.07 | 1.00 |
| | FR | CR | 0.15 | 0.07 | 0.04 |
| | | CG | 0.15 | 0.07 | 0.04 |
| | CG | CR | 0.00 | 0.07 | 1.00 |
| | | FR | -0.15 | 0.07 | 0.04 |
| 27 tense (past continuous tense) | CR | FR | -0.06 | 0.11 | 0.59 |
| | | CG | -0.06 | 0.11 | 0.59 |
| | FR | CR | 0.06 | 0.11 | 0.59 |
| | | CG | 0.00 | 0.11 | 1.00 |
| | CG | CR | 0.06 | 0.11 | 0.59 |
| | | FR | 0.00 | 0.11 | 1.00 |
| 28 non-finite verb | CR | FR | 0.12 | 0.11 | 0.28 |
| | | CG | 0.00 | 0.11 | 1.00 |
| | FR | CR | -0.12 | 0.11 | 0.28 |
| | | CG | -0.12 | 0.11 | 0.28 |
| | CG | CR | 0.00 | 0.11 | 1.00 |
| | | FR | 0.12 | 0.11 | 0.28 |
| 29 attributive clause | CR | FR | 0.15 | 0.11 | 0.19 |
| | | CG | 0.15 | 0.11 | 0.19 |
| | FR | CR | -0.15 | 0.11 | 0.19 |
| | | CG | 0.00 | 0.11 | 1.00 |
| | CG | CR | -0.15 | 0.11 | 0.19 |
| | | FR | 0.00 | 0.11 | 1.00 |
| 30 complex sentence (object clause) | CR | FR | -0.03 | 0.02 | 0.22 |
| | | CG | -0.03 | 0.02 | 0.22 |
| | FR | CR | 0.03 | 0.02 | 0.22 |
| | | CG | 0.00 | 0.02 | 1.00 |
| | CG | CR | 0.03 | 0.02 | 0.22 |
| | | FR | 0.00 | 0.02 | 1.00 |

| | | | | | |
|--|----|----|-------|------|------|
| 31 complex sentence (subject clause) | CR | FR | -0.09 | 0.04 | 0.03 |
| | | CG | -0.09 | 0.04 | 0.03 |
| | FR | CR | 0.09 | 0.04 | 0.03 |
| | | CG | 0.00 | 0.04 | 1.00 |
| | CG | CR | 0.09 | 0.04 | 0.03 |
| | | FR | 0.00 | 0.04 | 1.00 |
| 32 tense (simple past tense) | CR | FR | -0.06 | 0.11 | 0.59 |
| | | CG | 0.12 | 0.11 | 0.28 |
| | FR | CR | 0.06 | 0.11 | 0.59 |
| | | CG | 0.18 | 0.11 | 0.11 |
| | CG | CR | -0.12 | 0.11 | 0.28 |
| | | FR | -0.18 | 0.11 | 0.11 |
| 33 attributive clause | CR | FR | -0.09 | 0.08 | 0.26 |
| | | CG | 0.00 | 0.08 | 1.00 |
| | FR | CR | 0.09 | 0.08 | 0.26 |
| | | CG | 0.09 | 0.08 | 0.26 |
| | CG | CR | 0.00 | 0.08 | 1.00 |
| | | FR | -0.09 | 0.08 | 0.26 |
| 34 participle as attributive | CR | FR | -0.06 | 0.11 | 0.57 |
| | | CG | 0.06 | 0.11 | 0.57 |
| | FR | CR | 0.06 | 0.11 | 0.57 |
| | | CG | 0.12 | 0.11 | 0.26 |
| | CG | CR | -0.06 | 0.11 | 0.57 |
| | | FR | -0.12 | 0.11 | 0.26 |
| 35 adverbial clause | CR | FR | 0.00 | 0.10 | 1.00 |
| | | CG | 0.03 | 0.10 | 0.76 |
| | FR | CR | 0.00 | 0.10 | 1.00 |
| | | CG | 0.03 | 0.10 | 0.76 |
| | CG | CR | -0.03 | 0.10 | 0.76 |
| | | FR | -0.03 | 0.10 | 0.76 |
| 36 subject-verb agreement | CR | FR | -0.06 | 0.10 | 0.53 |
| | | CG | -0.03 | 0.10 | 0.75 |
| | FR | CR | 0.06 | 0.10 | 0.53 |
| | | CG | 0.03 | 0.10 | 0.75 |
| | CG | CR | 0.03 | 0.10 | 0.75 |
| | | FR | -0.03 | 0.10 | 0.75 |
| 37 gerund as subject | CR | FR | -0.03 | 0.09 | 0.74 |
| | | CG | 0.15 | 0.09 | 0.09 |
| | FR | CR | 0.03 | 0.09 | 0.74 |
| | | CG | 0.18 | 0.09 | 0.05 |
| | CG | CR | -0.15 | 0.09 | 0.09 |
| | | FR | -0.18 | 0.09 | 0.05 |
| 38 attributive clause | CR | FR | -0.09 | 0.07 | 0.22 |
| | | CG | 0.03 | 0.07 | 0.68 |
| | FR | CR | 0.09 | 0.07 | 0.22 |
| | | CG | 0.12 | 0.07 | 0.11 |
| | CG | CR | -0.03 | 0.07 | 0.68 |
| | | FR | -0.12 | 0.07 | 0.11 |
| 39 attributive clause | CR | FR | 0.03 | 0.12 | 0.80 |
| | | CG | 0.09 | 0.12 | 0.46 |
| | FR | CR | -0.03 | 0.12 | 0.80 |
| | | CG | 0.06 | 0.12 | 0.62 |
| | CG | CR | -0.09 | 0.12 | 0.46 |
| | | FR | -0.06 | 0.12 | 0.62 |
| 40 inversion | CR | FR | -0.09 | 0.10 | 0.35 |
| | | CG | 0.03 | 0.10 | 0.76 |
| | FR | CR | 0.09 | 0.10 | 0.35 |
| | | CG | 0.12 | 0.10 | 0.22 |
| | CG | CR | -0.03 | 0.10 | 0.76 |
| | | FR | -0.12 | 0.10 | 0.22 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

any absolute advantage. The Free Reading group was not leading on every item; similarly, the Control group did drop behind on every item either. For example, the Complementary Reading group was the best on Items 29 and 39, which both tested the attributive clause, and Item 25, which tested the gerund. The Free Reading group was leading on Item 32, simple past tense, Item 33, attributive clause and Item 34, participle as attributive. The Control group outperformed the Complementary Reading group on Item 27, simple past tense, Item 31, complex sentence and Item 36, subject-verb agreement. More interestingly, the Complementary Reading group and the Control group outperformed the Free Reading group on Item 28, non-finite verb.

In the previous discussion, it could not be concluded that extensive reading is effective in terms of incidental grammar acquisition. In this discussion, it cannot be concluded that extensive reading strengthens grammar rules that have been explicitly taught in the classroom. The detailed analysis on the three groups' performance on individual items has shown that no group has achieved absolute advantage on a specific grammar item, i.e., attributive clause, complex sentences or non-finite verb. Based on this, it is assumed that extensive reading did improve grammar competence; however, both the role of extensive reading in incidental grammar acquisition and the ways how grammar rules that were explicitly taught in classroom are consolidated still needs further study. The two experimental groups showed no clear advantage for either explicitly or implicitly taught grammar rules, even though they outperformed the Control group in the grammar section in the post-test. An interesting question that arises is: What did the two experimental groups actually acquire in their extensive reading?

Discussion 3: What did the participants actually acquire in their extensive reading?

Certain assumptions about first language acquisition propose it that children arrive at an internalized system, a grammar, on the basis of linguistic input. The acquisition task can be schematized in (i):

(i) Input →

| |
|---------|
| Grammar |
|---------|

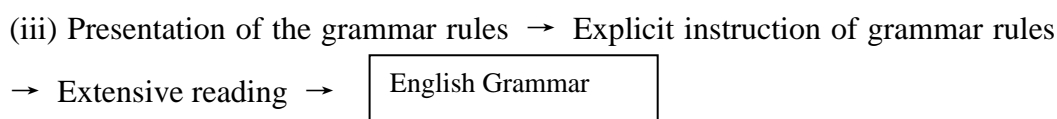
The L2 learners' acquisition task is very much like that of the L1 learner. L2 learners are also faced with the same problem of interpreting the input data, of establishing a system which will make sense of data, and which will enable them to understand and produce structures in their second language. Thus, their task can be conceived of as in (ii), equivalent to the L1 acquisition task:



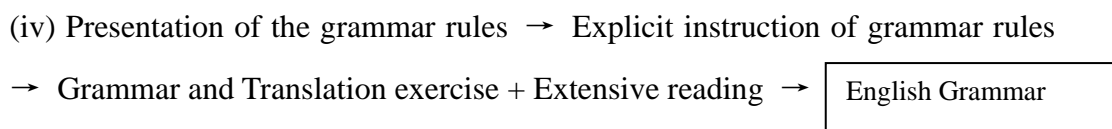
(Barker & McCarthy, 1981, p. 37).

However, the participants' task in this study bears no resemblance to that of the L1 and L2 learners. The Grammar-Translation Method has been criticized for providing no or little comprehensible input for acquisition.

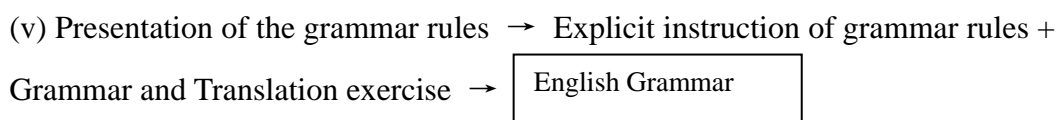
As schematized as (iii), the acquisition task of the Complementary Reading group is:



The acquisition task of the Free Reading group can be schematized as (iv):



The acquisition task of the Control group can be schematized as (v):



There were differences in the three schemata, but all of them showed that extensive reading was not the sole venue for grammar competence. Both presentation of the grammar rules and explicit instruction of the grammar rules are integral components in the process by which participants acquire grammar competence. Based on the test results alone, we are unable to identify the proportion that extensive reading accounts for in the development of the participants' grammar competence. Thus two areas of inquiry deserve further discussion: 1. Does extensive reading provide favorable conditions for grammar acquisition? 2. What is the role of reading quantity?

Discussion 4: Does extensive reading provide favorable conditions for grammar acquisition?

The focus of debate of this discussion is that graded readers are texts written by means of lexical, structural, and information control (Hedge, 1985; Hill, 2008). One characteristic of graded readers is that they strictly control grammar complexity for different levels. One criticism of graded readers is that they do not expose the students to a comprehensive range of English grammar constructions, as noted earlier in Chapter 2.

Concerning this debate, we should however be reminded that decoding difficult text is a slow, arduous business requiring, often at lower levels of proficiency, recoding into the L1 (Cohen, 1995). Because reading is slow, not much text is read, and there is not sufficient exposure to language input (Parry, 1991). In this sense, graded readers provide optimal comprehensible input for EFL learners.

A more direct method is to revisit the grammar syllabus for the participants in this program. Table 4.31 below presents the OBW syllabus and levels, and the grammar syllabus in the Shanghai English Curriculum Standards for Primary and Junior Middle School (2004) has been presented in Table 4.28. The comparison between the two syllabuses shows that the OBW graded readers are able to expose the readers to all the grammar rules they have known. In other words, the participants in this study were able to encounter all the grammar rules that they previously knew and that they were being taught in class. Although controlled in terms of grammar, graded readers are not less capable of providing optimal comprehensible input for grammar acquisition.

At this stage it cannot be concluded that extensive reading by means of graded readers provides optimal comprehensible input for grammar acquisition, but the fact that the Complementary Reading group read approximately five times as much as the Free Reading group but made smaller progress reminds us that the precise role that reading quantity plays in grammar acquisition deserves further discussion.

Table 4. 31 Oxford Bookworm grammar syllabus and levels

| Grammar structures | | Levels | | | | |
|--------------------------------------|---------------------------|---------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| simple present tense | Starter, 250 headwords | Level 1, 400 headwords | Level 2, 700 headwords | Level 3, 1000 headwords | Level 4, 1400 headwords | Level 5, 1800 headwords |
| present continuous tense | | | | | | |
| imperative | | | | | | |
| can/cannot, must | | | | | | |
| going to (future) | | | | | | |
| simple gerunds | | | | | | |
| simple past tense | Level 1, 400 headwords | Level 2, 700 headwords | Level 3, 1000 headwords | Level 4, 1400 headwords | Level 5, 1800 headwords | |
| present perfect tense | | | | | | |
| will (future) | | | | | | |
| (don't) have to, must not, could | | | | | | |
| comparison of adjectives | | | | | | |
| simple time clauses | | | | | | |
| past continuous tense | Level 2, 700 headwords | Level 3, 1000 headwords | Level 4, 1400 headwords | Level 5, 1800 headwords | Level 6, 2500 headwords | |
| tag questions | | | | | | |
| ask/tell infinitive | | | | | | |
| should, may | | | | | | |
| present perfect continuous | | | | | | |
| used to | | | | | | |
| past perfect tense | Level 3, 1000 headwords | Level 4, 1400 headwords | Level 5, 1800 headwords | Level 6, 2500 headwords | | |
| causative | | | | | | |
| relative clauses | | | | | | |
| indirect statements | | | | | | |
| past perfect continuous tense | | | | | | |
| passive (simple forms) | | | | | | |
| would conditional clauses | Level 4, 1400 headwords | Level 5, 1800 headwords | Level 6, 2500 headwords | | | |
| indirect questions | | | | | | |
| relatives with where/when | | | | | | |
| clauses of purpose, reason, contrast | | | | | | |
| gerunds after preposition/phrases | | | | | | |
| future continuous | | | | | | |
| future perfect | Level 5, 1800 headwords | Level 6, 2500 headwords | | | | |
| passive (modals, continuous forms) | | | | | | |
| would have conditional clauses | | | | | | |
| modal + perfect infinitive | | | | | | |
| so/such...that result clauses | | | | | | |
| passive (infinitives, gerunds) | | | | | | |
| advanced modal meanings | Level 6, 2500 headwords | | | | | |
| clauses of concession, condition | | | | | | |

(Retrieve

from: http://www.oup-bookworms.com/downloads/pdf/oxford_bookworms/syllabus.pdf)

Conclusion

In this section, the significance of acquiring grammar through extensive reading and the value of learning grammar through explicit instruction were examined. From the analysis and discussion above, it can be concluded that there is a positive relationship between extensive reading and grammar acquisition. The post-test result provided evidence in favor of extensive reading. The Free Reading group performed the best among the three. The Complementary Reading

group did not do grammar exercises after class but it did seem to be slightly better at grammar than the Control group although there is no significant main effect for the treatments as a factor in the vocabulary section in the post-test (CR=15.88, FR=16.64, CG=15.03, $p=0.07$).

A possible future research question could examine on which aspect of grammar extensive reading is more likely to exert influence. Is it sentence structure, sentence pattern or something else? Grammar is a very comprehensive concept, and some detailed classification is needed before it can be stated with certainty that extensive reading does positively influence grammar and in which part of grammar it has a bigger role to play.

A closer look at the syllabuses and examination papers suggest that success in the pre-, mid- and post-tests is possible with limited knowledge and understanding of grammatical structures or systems. The Free Reading group and the Control group's repeated grammar and translation exercises after school as well as standard "grammatical" paradigm and latter's to be adapted by students according to context appear to be successful strategies in improving examination results.

4.5 Extensive reading and receptive skills

4.5.1 Extensive reading and reading comprehension

In the National Matriculation English Test (Shanghai version), the text comprehension section, which accounts for 33.3% in this highest-stakes English exam, has two complementary components: cloze and reading comprehension (see Section 1.4 for more information on this test). This design of text comprehension reflects that the test designers on one hand wish to isolate one aspect of reading ability or one aspect of language; on the other hand, they expect a global overview of the test takers' ability that whether the test takers are able to comprehend the texts satisfactorily.

As shown in Table 4.32, there is no significant main effect for the treatments as a factor in cloze and reading comprehension in the pre-test. It suggested that the three groups joined this extensive reading program with approximately equal

English language reading proficiency. The situation changed in the mid-test. There is still no significant main effect for the treatments as a factor in cloze, but there is a main effect for the treatments as a factor in reading ($F=3.69$, $p=0.03$). Judging from the mean scores, the Free Reading group performed the best among the three in reading with a 0.03 point advantage to the Complementary Reading group and a 1.33 point advantage to the Control group ($CR=12.21$, $FR=12.24$, $CG=10.88$). In the post-test, the gap further widened. There is a significant main effect for the treatments as a factor in both cloze and reading sections. In terms of mean score, in the cloze section whose total score was ten points, the Free Reading group obtained advantage by 0.24 point to the Complementary Reading group and 1.46 point to the Control group ($CR=6.55$, $FR=6.79$, $CG=5.33$). In the reading comprehension section, the Free Reading group took lead by a 0.03 point advantage to the Complementary Reading group and a 1.34 point advantage to the Control group ($CR=11.55$, $FR=11.58$, $CG=10.24$).

Table 4. 32 Descriptive statistics and results of ANOVA from the cloze and reading comprehension sections in the pre-, mid- and post-tests

| Tests | Sections | Group | N | Mean | SD | F | Sig. |
|-----------|-----------|-------|----|-------|------|------|------|
| Pre-test | Cloze | CR | 33 | 14.18 | 2.51 | 0.75 | 0.48 |
| | | FR | 33 | 13.52 | 2.59 | | |
| | | CG | 33 | 13.58 | 2.24 | | |
| | Reading 1 | CR | 33 | 12.48 | 2.65 | 1.01 | 0.37 |
| | | FR | 33 | 12.55 | 3.40 | | |
| | | CG | 33 | 11.58 | 3.23 | | |
| | Reading2 | CR | 33 | 5.24 | 1.56 | 0.35 | 0.71 |
| | | FR | 33 | 5.18 | 1.57 | | |
| | | CG | 33 | 5.48 | 1.56 | | |
| Mid-test | Cloze | CR | 33 | 7.79 | 1.27 | 0.58 | 0.56 |
| | | FR | 33 | 7.58 | 1.60 | | |
| | | CG | 33 | 7.36 | 1.87 | | |
| | Reading | CR | 33 | 12.21 | 2.51 | 3.69 | 0.03 |
| | | FR | 33 | 12.24 | 2.06 | | |
| | | CG | 33 | 10.88 | 2.39 | | |
| Post-test | Cloze | CR | 33 | 6.55 | 1.35 | 9.20 | 0.00 |
| | | FR | 33 | 6.79 | 1.63 | | |
| | | CG | 33 | 5.33 | 1.43 | | |
| | Reading | CR | 33 | 11.55 | 1.95 | 5.33 | 0.01 |
| | | FR | 33 | 11.58 | 2.08 | | |
| | | CG | 33 | 10.24 | 1.62 | | |

* The mean difference is significant at the 0.05 level.

Note 1: The results of Levene tests show there is no deviation from homogeneity of variance.

Note 2: Reading 1 was in the form of multiple-choice and Reading 2 was short questions and answers.

As illustrated in Table 4.33, all the significant differences were between the two

experimental groups and the Control group. There was no statistically significant difference between the two experimental groups in either cloze or reading sections.

Cloze tests are typically constructed by deleting from selected texts and simply requiring the test takers to restore the words that have been deleted. What an individual cloze test measures depends on which individual words are deleted. Depending on which individual words are deleted, the ten items in the cloze section in the post-test can be categorized into three groups proposed by Alderson (2000): items measuring sensitivity to lexical semantics, items constrained by short-range discourse and items constrained by long-range discourse.

Table 4. 33The results of multiple comparisons by the Fisher's LSD test in the cloze and reading comprehension sections in the post-test

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| Cloze | CR | FR | -0.24 | 0.36 | 0.51 |
| | | CG | 1.21 | 0.36 | 0.00 |
| | FR | CR | 0.24 | 0.36 | 0.51 |
| | | CG | 1.45 | 0.36 | 0.00 |
| | CG | CR | -1.21 | 0.36 | 0.00 |
| | | FR | -1.45 | 0.36 | 0.00 |
| Reading | CR | FR | -0.03 | 0.47 | 0.95 |
| | | CG | 1.30 | 0.47 | 0.01 |
| | FR | CR | 0.03 | 0.47 | 0.95 |
| | | CG | 1.33 | 0.47 | 0.01 |
| | CG | CR | -1.30 | 0.47 | 0.01 |
| | | FR | -1.33 | 0.47 | 0.01 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Items 51, 55, 56, and 57 appear to measure sensitivity to semantics. For example, on Item 51, the test takers needed to locate a word among “ways, trends, use, disadvantages” which describes “build our bodies, prevent us from gaining weights, and keep us healthy”. Items 50 and 54 are not constrained by long-range discourse but by the immediately adjacent sentence. The answer to Item 50 is decided by the immediately adjacent sentence “people do at times when they are not working”, which implies that sports and games are “unimportant” things. Item 54 requires an item from four possible conjunctions by the need for coherence with the immediately adjacent sentence. Item 58 and 59 can be said to

be constrained by larger units of discourse than the sentence.

To summarize, almost half of the items, i.e. Items 51, 55, 56 and 57, measure sensitivity to lexical semantics while two of them, i.e. Items 50 and 54 are constrained by immediately adjacent sentences, and Items 58 and 59 measure sensitivity to the topic of the text, though not necessarily to the meaning of the whole passage.

As shown in Table 4.34 below, the three groups performed significantly differently on Items 53 and 54. As introduced previously, Item 53 measures sensitivity to the process of ping pong, but not necessarily to the meaning of the whole paragraph or passage. On this item, there is a significant main effect for the treatments as a factor ($CR=1.00$, $FR=1.00$, $CG=0.91$, $p=0.05$). Item 54, which requires an item from four possible conjunctions, is constrained by the need for coherence with the immediately adjacent sentence. The two experimental groups showed significantly better ability to manage the coherence between adjacent sentences by scoring nearly twice as high as the Control Group ($CR=0.64$, $FR=0.61$, $CG=0.33$, $p=0.02$). In addition to the main effect on Items 53 and 54, a practical difference between the three groups is found on Item 58, although the difference is not statistically significant ($F=2.49$, $p=0.09$). On Items 58, which measures lexical sensitivity to the topic of the text, the two experimental groups scored nearly 40% higher than the Control group ($CR=0.73$, $FR=0.79$, $CG=0.55$). Synthesizing the three groups' performance on these three items, it is suggested that there is no significant main effect for the treatments as a factor in answering items measuring sensitivity to lexical semantics. Conversely, the difference among the three groups lies in their ability to manage the discourse of the text, namely the flow of the message.

Table 4.35 below confirms that there is no statistically significance between the two experimental groups on Item 53, 54 and 58. In contrast, both experimental groups outperformed the Control group on these three items.

Table 4. 34 Descriptive statistics and results of ANOVA from the cloze section in the post-test

| Item | Group | N | Mean | SD | Std. Error | F | Sig. |
|------|-------|----|------|------|------------|------|------|
| 50 | CR | 33 | 0.58 | 0.50 | 0.09 | 2.12 | 0.13 |
| | FR | 33 | 0.79 | 0.42 | 0.07 | | |
| | CG | 33 | 0.76 | 0.44 | 0.08 | | |
| 51 | CR | 33 | 0.85 | 0.36 | 0.06 | 0.23 | 0.80 |
| | FR | 33 | 0.88 | 0.33 | 0.06 | | |
| | CG | 33 | 0.82 | 0.39 | 0.07 | | |
| 52 | CR | 33 | 0.58 | 0.50 | 0.09 | 2.13 | 0.12 |
| | FR | 33 | 0.52 | 0.51 | 0.09 | | |
| | CG | 33 | 0.33 | 0.48 | 0.08 | | |
| 53 | CR | 33 | 1.00 | 0.00 | 0.00 | 3.20 | 0.05 |
| | FR | 33 | 1.00 | 0.00 | 0.00 | | |
| | CG | 33 | 0.91 | 0.29 | 0.05 | | |
| 54 | CR | 33 | 0.64 | 0.49 | 0.09 | 3.86 | 0.02 |
| | FR | 33 | 0.61 | 0.50 | 0.09 | | |
| | CG | 33 | 0.33 | 0.48 | 0.08 | | |
| 55 | CR | 33 | 0.91 | 0.29 | 0.05 | 0.70 | 0.50 |
| | FR | 33 | 0.82 | 0.39 | 0.07 | | |
| | CG | 33 | 0.82 | 0.39 | 0.07 | | |
| 56 | CR | 33 | 0.39 | 0.50 | 0.09 | 1.71 | 0.19 |
| | FR | 33 | 0.45 | 0.51 | 0.09 | | |
| | CG | 33 | 0.24 | 0.44 | 0.08 | | |
| 57 | CR | 33 | 0.55 | 0.51 | 0.09 | 0.04 | 0.96 |
| | FR | 33 | 0.52 | 0.51 | 0.09 | | |
| | CG | 33 | 0.52 | 0.51 | 0.09 | | |
| 58 | CR | 33 | 0.73 | 0.45 | 0.08 | 2.49 | 0.09 |
| | FR | 33 | 0.79 | 0.42 | 0.07 | | |
| | CG | 33 | 0.55 | 0.51 | 0.09 | | |
| 59 | CR | 33 | 0.33 | 0.48 | 0.08 | 1.85 | 0.16 |
| | FR | 33 | 0.33 | 0.48 | 0.08 | | |
| | CG | 33 | 0.15 | 0.36 | 0.06 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

The data analysis in the cloze section suggests that the two experimental groups were superior in handling the text at discourse level. Subsequently, their performance in the reading comprehension section will be examined.

The test designers aim to gain a much more general idea of how well the students could comprehend the texts. According to Davis (1968), there are eight levels of understanding as follows:

1. recalling word meanings
2. drawing inferences about the meaning of a word in context (Item 62)
3. finding answers to questions answered explicitly or in paraphrase (Items 60, 61, 63, 66)

4. weaving together ideas in the content
5. drawing inferences from the content (Items 64, 65, 68, 69)
6. recognizing a writer's purpose, attitude, tone and mood (Items 67, 70)
7. identifying a writer's technique
8. following the structure of a passage

(cited in Alderson, 2000, p. 9-10)

Table 4. 35 The results of multiple comparisons by the Fisher's LSD test in the cloze section in the post-test

| Item | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|------|-----------|-----------|-----------------------|------------|------|
| 53 | CR | FR | 0.00 | 0.04 | 1.00 |
| | | CG | 0.09 | 0.04 | 0.03 |
| | FR | CR | 0.00 | 0.04 | 1.00 |
| | | CG | 0.09 | 0.04 | 0.03 |
| | CG | CR | -0.09 | 0.04 | 0.03 |
| | | FR | -0.09 | 0.04 | 0.03 |
| 54 | CR | FR | 0.03 | 0.12 | 0.80 |
| | | CG | 0.30 | 0.12 | 0.01 |
| | FR | CR | -0.03 | 0.12 | 0.80 |
| | | CG | 0.27 | 0.12 | 0.03 |
| | CG | CR | -0.30 | 0.12 | 0.01 |
| | | FR | -0.27 | 0.12 | 0.03 |
| 58 | CR | FR | -0.06 | 0.11 | 0.59 |
| | | CG | 0.18 | 0.11 | 0.11 |
| | FR | CR | 0.06 | 0.11 | 0.59 |
| | | CG | 0.24 | 0.11 | 0.03 |
| | CG | CR | -0.18 | 0.11 | 0.11 |
| | | FR | -0.24 | 0.11 | 0.03 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

The eleven items in the reading comprehension section in the post-test did not cover all eight levels defined by Davis. Among the eleven items, Item 62 was devoted to Level 2, Items 60, 61, 63 and 66 were devoted to Level 3, Items 64, 65, 68 and 69 belonged to Level 5, and Items 67 and 70 were at Level 6

As can be seen from Table 4.36, the three groups demonstrated no statically significant difference in their ability to make sense of a lexical item by using contextual clues, namely Item 62, and scan for specific information, namely Items 60, 61, 63 and 66. The statistically significant difference among the three groups appeared on Item 67 and 68, and a near significant difference on Item 66. In other words, statistically significant difference was found on some of the items

(though not all) which measure more global reading skills such as inferencing skills and recognition of writers' stance.

Table 4. 36 Descriptive statistics and results of ANOVA from the reading comprehension section in the post-test

| Item | Group | N | Mean | SD | Std. Error | F | Sig. |
|------|-------|----|------|------|------------|------|------|
| 60 | CR | 33 | 0.88 | 0.33 | 0.06 | 0.23 | 0.80 |
| | FR | 33 | 0.82 | 0.39 | 0.07 | | |
| | CG | 33 | 0.85 | 0.36 | 0.06 | | |
| 61 | CR | 33 | 0.79 | 0.42 | 0.07 | 2.19 | 0.12 |
| | FR | 33 | 0.88 | 0.33 | 0.06 | | |
| | CG | 33 | 0.67 | 0.48 | 0.08 | | |
| 62 | CR | 33 | 0.73 | 0.45 | 0.08 | 0.05 | 0.95 |
| | FR | 33 | 0.76 | 0.44 | 0.08 | | |
| | CG | 33 | 0.76 | 0.44 | 0.08 | | |
| 63 | CR | 33 | 0.21 | 0.42 | 0.07 | 0.06 | 0.94 |
| | FR | 33 | 0.24 | 0.44 | 0.08 | | |
| | CG | 33 | 0.21 | 0.42 | 0.07 | | |
| 64 | CR | 33 | 0.82 | 0.39 | 0.07 | 2.00 | 0.14 |
| | FR | 33 | 0.82 | 0.39 | 0.07 | | |
| | CG | 33 | 0.64 | 0.49 | 0.09 | | |
| 65 | CR | 33 | 0.94 | 0.24 | 0.04 | 0.44 | 0.65 |
| | FR | 33 | 0.88 | 0.33 | 0.06 | | |
| | CG | 33 | 0.88 | 0.33 | 0.06 | | |
| 66 | CR | 33 | 0.67 | 0.48 | 0.08 | 2.73 | 0.07 |
| | FR | 33 | 0.67 | 0.48 | 0.08 | | |
| | CG | 33 | 0.42 | 0.50 | 0.09 | | |
| 67 | CR | 33 | 0.85 | 0.36 | 0.06 | 3.75 | 0.03 |
| | FR | 33 | 0.70 | 0.47 | 0.08 | | |
| | CG | 33 | 0.55 | 0.51 | 0.09 | | |
| 68 | CR | 33 | 0.64 | 0.49 | 0.09 | 3.86 | 0.02 |
| | FR | 33 | 0.61 | 0.50 | 0.09 | | |
| | CG | 33 | 0.33 | 0.48 | 0.08 | | |
| 69 | CR | 33 | 0.55 | 0.51 | 0.09 | 0.48 | 0.62 |
| | FR | 33 | 0.61 | 0.50 | 0.09 | | |
| | CG | 33 | 0.48 | 0.51 | 0.09 | | |
| 70 | CR | 33 | 0.24 | 0.44 | 0.08 | 0.84 | 0.44 |
| | FR | 33 | 0.21 | 0.42 | 0.07 | | |
| | CG | 33 | 0.12 | 0.33 | 0.06 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

The Fishers' LSD test was conducted on Items 66, 67 and 68. As shown in Table 4.37 below, the two experimental groups significantly outperformed the Control group on Items 67 and 68. As for Item 66, both experimental groups significantly outperformed the Control group ($p=0.05$) although the descriptive statistic presented in Table 4.36 suggests that there was no significant main effect for the treatments as a factor on this item.

The statistically significant difference on Item 66, 67 and 68 suggests that the two experimental groups demonstrated higher ability on three of the four levels of understanding tested, from scanning for specific information to making sense

of a lexical item by using contextual clues and recognizing the writer's stance. It indicates that extensive reading enables the students to be better at reading the lines, and reading beyond the lines as well.

Table 4. 37 The results of multiple comparisons by the Fisher's LSD test on Items 66, 67 and 68 in the reading comprehension section in the post-test

| Item | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|------|-----------|-----------|-----------------------|------------|------|
| 66 | CR | FR | 0.00 | 0.12 | 1.00 |
| | | CG | 0.24 | 0.12 | 0.05 |
| | FR | CR | 0.00 | 0.12 | 1.00 |
| | | CG | 0.24 | 0.12 | 0.05 |
| | CG | CR | -0.24 | 0.12 | 0.05 |
| | | FR | -0.24 | 0.12 | 0.05 |
| 67 | CR | FR | 0.15 | 0.11 | 0.17 |
| | | CG | 0.30 | 0.11 | 0.01 |
| | FR | CR | -0.15 | 0.11 | 0.17 |
| | | CG | 0.15 | 0.11 | 0.17 |
| | CG | CR | -0.30 | 0.11 | 0.01 |
| | | FR | -0.15 | 0.11 | 0.17 |
| 68 | CR | FR | 0.03 | 0.12 | 0.80 |
| | | CG | 0.30 | 0.12 | 0.01 |
| | FR | CR | -0.03 | 0.12 | 0.80 |
| | | CG | 0.27 | 0.12 | 0.03 |
| | CG | CR | -0.30 | 0.12 | 0.01 |
| | | FR | -0.27 | 0.12 | 0.03 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Another point, which is a by-product of this discussion, is that the processing of linguistic information is central to reading comprehension (Perfetti, et al., 2005). The levels of understanding of the texts, which were measured by sixteen multiple-choice items, are based on the fact that the readers comprehend the texts. It is not possible to make sense of a lexical item by using contextual clues, nor predict the flow of information without recognizing the words to be read and the key information indicating the writer's stance. The two experimental groups' superior performance in this sub-section, from one perspective, demonstrates that extensive reading is a contributing factor to reading proficiency; from another perspective, it confirms that, the two experimental groups showed richer linguistic knowledge, i.e., morphological, syntactic and semantic knowledge, in this subsection.

The last five items, Items 71 to 75, measures differently from Items 60 to 70. Recent language tests have experimented with a number of techniques for

assessing reading. One of the techniques is multiple matching, in which two sets of stimuli are to be matched with each other, for instance, matching a heading for its corresponding paragraph (Alderson, 2000). In the post-test, the students would need to match the six suggested headings to fit the five paragraphs. These five paragraphs are not isolated or independent from each other but from the same passage. Although each of them is intended to be matched with a different heading, they are actually about the same topic. In the post-test, the five paragraphs talk about English language from five different but relevant perspectives. To match the two sets of stimuli correctly, the students would need to comprehend each paragraph and the whole passage as well.

As shown in Table 4.38, there is no significant main effect for the treatments as a factor in the title-matching section, which indicates that the three groups show no significantly different ability in comprehending the passage.

Table 4. 38 Descriptive statistics and results of ANOVA from the title-matching subsection in the post-test

| Item | Group | N | Mean | SD | Std. Error | F | Sig. |
|------|-------|----|------|------|------------|------|------|
| 71 | CR | 33 | 0.85 | 0.36 | 0.06 | 0.77 | 0.47 |
| | FR | 33 | 0.91 | 0.29 | 0.05 | | |
| | CG | 33 | 0.94 | 0.24 | 0.04 | | |
| 72 | CR | 33 | 0.94 | 0.24 | 0.04 | 2.06 | 0.13 |
| | FR | 33 | 1.00 | 0.00 | 0.00 | | |
| | CG | 33 | 1.00 | 0.00 | 0.00 | | |
| 73 | CR | 33 | 0.91 | 0.29 | 0.05 | 0.13 | 0.88 |
| | FR | 33 | 0.91 | 0.29 | 0.05 | | |
| | CG | 33 | 0.94 | 0.24 | 0.04 | | |
| 74 | CR | 33 | 0.67 | 0.48 | 0.08 | 0.37 | 0.69 |
| | FR | 33 | 0.58 | 0.50 | 0.09 | | |
| | CG | 33 | 0.58 | 0.50 | 0.09 | | |
| 75 | CR | 33 | 0.88 | 0.33 | 0.06 | 0.53 | 0.59 |
| | FR | 33 | 0.94 | 0.24 | 0.04 | | |
| | CG | 33 | 0.94 | 0.24 | 0.04 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

To summarize, in this section, the relationship between extensive reading and reading proficiency has been discussed. On the extensive reading and reading proficiency relationship, the first research area focuses on readers' improvement in reading strategies and skills. In Constantino's (1995) study, twenty ESL learners read for pleasure for twenty minutes in each class. At the very beginning, they read in a word-by-word, sentence-by-sentence manner. When encountering

an unknown lexical item, they turned to a bilingual dictionary; when having difficulty building clause-level meaning, they tried to make sense of the sentence by analyzing its grammatical structure. With extensive reading going on, such practice decreased dramatically or vanished. They become more fluent readers by employing more reading skills and strategies. Liem (2005) utilized extensive reading to enhance students' reading ability in Vietnam. Data collected from the post-program questionnaire and interview showed that extensive reading resulted in the participants radical rethinking their meta-cognitive strategies acquired before, and apply these strategies more effectively in their future reading. Ramaiah (1994) noticed students' shift from bottom-up to more top-down or interactive views of the reading process after the reading program.

The second research area primarily concentrates on students' reading ability and comprehension performance. For instance, Kutiper (1983) suggested that extensive reading was as effective as intensive reading in developing general reading ability among elementary and secondary school students towards reading. For another example, Krashen (1988) examined the relationship between the amount of free reading done and reading ability, as measured by tests of reading comprehension, arguing that free reading consistently related to success in reading comprehension.

The post-test results were consistent with the second research stream. As evidenced by the post-test results, there was a statistically significant difference between the two experimental groups and the control group in both cloze and reading comprehension sections, which indicated that their reading ability and comprehension performance were superior after their extensive reading experience.

4.5.2 Extensive reading and reading fluency

Because of its effect in language teaching and acquisition, extensive reading in a second and foreign language has become a motivating way of language teaching (Han, 2010). However, some researchers stress that students need to read for a real purpose, rather than focusing on extensive reading as a tool for teaching language (e.g., Hyland, 1990). The researchers' focus on the effect of extensive

reading on the improvement of reading ability is highly understandable because the ability to read both in L1 and L2 is not only important for academic achievement but also professional and self-development.

The research literature concerning extensive reading and reading proficiency can be categorized into four research areas. As introduced in the previous discussion section, the first research area about the relationship between extensive reading and reading ability focuses on students' improvement in reading strategies and skills, and the second research area primarily concentrates on the students' reading ability and comprehension performance. In addition to the above-noted two areas, the third area comes from researchers' attention to reading fluency. The development of reading fluency has taken a position of growing importance in L2 research and has emerged as a significant pedagogical issue (Gorsuch & Taguchi, 2009). Grabe (1991) pointed out that sustained silent reading "build fluency (automaticity), confidence, and appreciation of reading".

As noted earlier in this section, our discussion in the previous section, namely Section 4.5.1, addressed the relationship between extensive reading and reading comprehension, in which the participants' ability to manage short texts was examined. In this section, we are going to explore the relationship between extensive reading and EPER levels, in which the participants' ability to handle longer texts will be investigated.

In this discussion section, two areas of inquiry are to be addressed: (1) what the effect of extensive reading is on improving students' reading fluency, and (2) what the scope and significance of the EPER placement tests is in the present study.

The pre- and post-EPER test results have been detailed in Section 4.3.2. To facilitate our discussion, some of the tables in Section 4.3.2 will be repeated in this section. As shown in Table 4.17 below, the difference between the pre- and post-EPER tests showed the change in EPER levels of the whole sample. The absolute numbers of Level A and B participants increases sharply. The percentage of Level A students increased from 1% to 5%, five times as many as

that in the pre-test. The number of Level B students doubled, from 19% to 40%. In addition, the number of Level D students decreased by 50% and the Level E students disappeared. From the table, a general trend was observed that the overall EPER levels of the sample move onto a higher stage.

Table 4. 17 Pre- and post-EPER levels of the whole sample (repeated)

| EPER Level | | A | B | C | D | E | Total |
|------------|--------|------|-------|-------|-------|------|--------|
| Headwords | | 2200 | 1900 | 1600 | 1200 | 800 | / |
| Pre-test | Number | 1 | 19 | 41 | 36 | 2 | 99 |
| | % | 1.01 | 19.19 | 41.41 | 36.36 | 2.02 | 100.00 |
| Post-test | Number | 5 | 40 | 37 | 17 | / | 99 |
| | % | 5.05 | 40.40 | 37.37 | 17.17 | / | 100.00 |

Table 4.18, the Group*Progress cross tabulation, illustrates that ten students in the whole sample slipped back by one level. Thirty-eight students remained stable in terms of their EPER levels. Forty-three students, the largest group in the cross tabulation, moved forward by one level. Seven students made progress by two levels and one student advanced by three levels.

Table 4. 18 Group * Progress Cross tabulation of the EPER Placement Test of the whole sample (repeated)

| Progress in level | | -1 | 0 | 1 | 2 | 3 | Total |
|-------------------|-------|----|----|----|---|---|-------|
| Group | CR | 1 | 10 | 16 | 5 | 1 | 33 |
| | FR | 4 | 11 | 16 | 2 | 0 | 33 |
| | CG | 5 | 17 | 11 | 0 | 0 | 33 |
| | Total | 10 | 38 | 43 | 7 | 1 | 99 |

Specifically, in the Complementary Reading group, sixteen students progressed by one level, five students advanced by two levels, and one student improved by three levels. In the Free Reading group, sixteen students progressed by one level, and two students advanced by two levels,. No student in this group made progress by three levels or more. In the Control group, eleven students progressed by one level. No students advanced by two levels or more.

To summarize, the progress in the two experimental groups was impressive. In the Complementary Reading group, sixteen students advanced by one level, five students by two levels and one student by three levels. In all, 66.7% of the participants in the Complementary Reading group made progress in terms of EPER level. In the Free Reading group, sixteen students advanced by one level

and two students by two levels. In all, 54.5% of the participants in the Free Reading group made progress in terms of EPER level. In contrast to the two experimental groups, the progress made by the Control group in terms of EPER level was less noticeable. In the Control group, eleven students advanced by one level. Further, no participant among the eleven of them made progress by two levels or more. On the contrary, five participants in the Control group fell back by one level, which was the largest number among the three groups. The Control group's performance on the EPER test shows that after a whole year's studying in senior high school, the students in the Control group still could not read more difficult reading materials in English. The weakness of the traditional Grammar-Translation Method was revealed as ineffective in improving students' reading proficiency in terms of EPER level. The problem becomes more acute when the National Matriculation English Test (Shanghai version) committee has been increasing the weighting that reading accounts for in this exam from 20% to 33% from 2004 to 2008, as shown in Table 1.15 in Chapter One.

The Chi-square test of the Group * Progress Cross tabulation confirmed the researcher's observation (see Table 4.19 below). The test produced evidence of statistically significant differences between pre- and post-test in terms of students' EPER levels ($p=0.001$, Chi-square=18.26, $df=4$). The progress in terms of EPER level did not happen by chance or sample errors but was attributed to extensive reading.

Table 4. 19 Chi-Square Tests of the results of the pre- and post-EPER Placement Test (repeated)

| Chi-Square Tests | | | |
|--------------------|------------|-------|-----------------------|
| | Value | df | Asymp. Sig. (2-sided) |
| Pearson Chi-Square | 18.26 | 4 | 0.001 |
| Likelihood Ratio | 19.55 | 4 | 0.001 |
| N of Valid Cases | 198 | | |
| Symmetric Measures | | | |
| | | Value | Approx. Sig. |
| Nominal by Nominal | Phi | .304 | .001 |
| | Cramer's V | .304 | .001 |
| N of Valid Cases | | 198 | |

* The mean difference is significant at the 0.05 level.

The general statistically significant difference between the pre- and post-EPER

tests of the whole sample was observed and confirmed. At the next step, we are going to detect where the actual difference lies by examining the test results of individual groups separately. The Complementary Reading group comes first. As illustrated in Table 4.39 below, the change between the pre- and post-tests is easily visible, as this group has two Level A students in the post-test, and the majority of the students (57.58%) are at Level B or higher now, while only 15.16% of the Complementary Reading participants attained these levels in the pre-test.

Table 4. 39 Pre- and post-EPER levels of the Complementary Reading group

| EPER Levels | | A | B | C | D | E | Total |
|-------------|--------|------|-------|-------|-------|------|--------|
| Headwords | | 2200 | 1900 | 1600 | 1200 | 800 | / |
| Pre-test | Number | / | 5 | 13 | 14 | 1 | 33 |
| | % | / | 15.16 | 39.40 | 42.42 | 3.03 | 100.00 |
| Post-test | Number | 2 | 17 | 12 | 2 | / | 33 |
| | % | 6.06 | 51.52 | 36.36 | 6.06 | / | 100.00 |

By conducting the Chi-Square Test, evidence was found of statistically significant differences between the pre- and post-test in terms of EPER levels of the Complementary Reading group ($p < 0.001$, Chi-square=17.022, $df=4$). As evidenced by the Symmetric Measures, the strength of this relationship between extensive reading and EPER level was found to be relatively strong ($\Phi=0.51$) (see Table 4.40 below).

Table 4. 40 Chi-Square test of the results of the pre- and post-EPER Placement Test of the Complementary Reading Group

| | Value | df | Asymp. Sig. (2-sided) |
|---|------------|-------|-----------------------|
| Pearson Chi-Square | 17.022(a) | 4 | .002 |
| Likelihood Ratio | 19.334 | 4 | .001 |
| N of Valid Cases | 66 | | |
| a 4 cells (40.0%) have expected count less than 5. The minimum expected count is .50. | | | |
| Symmetric Measures | | | |
| | | Value | Approx. Sig. |
| Nominal by Nominal | Phi | 0.51 | 0.00 |
| | Cramer's V | 0.51 | 0.00 |
| N of Valid Cases | | 66 | |

When it comes to the Free Reading group, the change is also visible as evidenced by two new Level A students, more Level B students and fewer Level D students (see Table 4.41 below).

Table 4. 41 Pre and post-EPER levels of the Free Reading group

| EPER Levels | | A | B | C | D | E | Total |
|-------------|--------|------|-------|-------|-------|---|--------|
| Pre-test | Number | / | 9 | 13 | 11 | / | 33 |
| | % | / | 27.28 | 39.40 | 33.33 | / | 100.00 |
| Post-test | Number | 2 | 15 | 11 | 5 | / | 33 |
| | % | 6.06 | 45.46 | 33.33 | 15.15 | / | 100.00 |

No statistically significant difference is found between the pre- and post-test in terms of the free reading participants' EPER levels ($p=0.116$, Chi-square=5.917, $df=3$) (see Table 4.42 below). The Symmetric Measures suggests there is no relationship between extensive reading and EPER level of the Free Reading group ($\Phi=0.299$).

Table 4. 42 Chi-Square test of the results of pre- and post-EPER Placement Test of the Free Reading Group

| | Value | df | Asymp. Sig. (2-sided) |
|--------------------|----------|----|-----------------------|
| Pearson Chi-Square | 5.917(a) | 3 | .116 |
| Likelihood Ratio | 6.761 | 3 | .080 |
| N of Valid Cases | 66 | | |

a 2 cells (25.0%) have expected count less than 5. The minimum expected count is 1.00.

| Symmetric Measures | | | |
|--------------------|------------|-------|--------------|
| | | Value | Approx. Sig. |
| Nominal by Nominal | Phi | .299 | .116 |
| | Cramer's V | .299 | .116 |
| N of Valid Cases | | 66 | |

Compared with those of the two experimental groups, the statistic of the Control Group is stable (see Table 4.43 below). The proportion of Level A student remains the same, and those of the Level B and Level D students have hardly changed.

Table 4. 43Pre- and post- EPER levels of the Control Group

| EPER Levels | | A | B | C | D | E | Total |
|-------------|--------|------|-------|-------|-------|------|--------|
| Headwords | | 2200 | 1900 | 1600 | 1200 | 800 | / |
| Pre-test | Number | 1 | 5 | 15 | 11 | 1 | 33 |
| | % | 3.03 | 15.16 | 45.46 | 33.33 | 3.03 | 100.00 |
| Post-test | Number | 1 | 8 | 14 | 10 | / | 33 |
| | % | 3.03 | 24.24 | 42.42 | 30.30 | / | 100.00 |

As shown in Table 4.44, no evidence was found of any statistically significant difference between the pre- and post-test in terms of the students' EPER levels in the control group ($p=0.777$, Chi-square=1.774, $df=4$). The fact that the present finding is not statistically significant, therefore, could possibly reflect the fact

that the traditional Grammar-Translation Method did not result in students' progress in EPER level.

Table 4. 44 Chi-Square test of the results of the pre- and post-EPER Placement Test of the Control Group

| | Value | df | Asymp. Sig. (2-sided) |
|--------------------|-----------|----|-----------------------|
| Pearson Chi-Square | 1.774 (a) | 4 | .777 |
| Likelihood Ratio | 2.167 | 4 | .705 |
| N of Valid Cases | 66 | | |

a 4 cells (40.0%) have expected count less than 5. The minimum expected count is .50.

| Symmetric Measures | | | |
|--------------------|------------|-------|--------------|
| | | Value | Approx. Sig. |
| Nominal by Nominal | Phi | .164 | .777 |
| | Cramer's V | .164 | .777 |
| N of Valid Cases | | 66 | |

To summarize the analysis of the results of the EPER Placement Test B, it can be concluded that the Complementary Reading Treatment is effective in improving students' EPER levels, while the Free Reading Treatment is less effective, and the traditional Grammar-Translation Method is the weakest among the three.

Having presented the EPER test results, our next step is to interpret the test EPER results, i.e., the participants' progress in terms of EPER level in the framework of reading fluency. The significance of reading fluency is to allow readers to experience a much large amount of reading materials, the expand the breadth and depth of vocabulary knowledge beyond explicit instruction, to develop automatic word recognition skills, to read for additional learning, to build reading motivation, and in L2 university context, to read large amounts of materials that might be assigned every week (Grabe, 2009). Most importantly, fluency is one of the keys to English language learning outside classroom. Students who have some degree of reading fluency and who are motivated to develop fluency further will most be likely to be engaged in a continual English language learning environment.

According to Kuhn and Stahl (2003), the three central sub-processes of fluency are automaticity, accuracy and rate. Consistent with Kuhn and Stahl, Gorsuch and Taguchi (2009) point out that the components of reading fluency are "readers' rapidity, accuracy and automaticity in fulfilling word recognition tasks" (p. 249). Segalowitz (2000), one of the leading L2 researchers in fluency and automaticity

processing, defines fluency as fluidity, speed and accuracy. It can be seen that the consensus on defining reading fluency is that the three core components are rate, automaticity and accuracy.

With reference to the first component in reading fluency, rate (rapidity or fluidity), findings from the participants' reading records demonstrated that the experimental participants achieved substantial gains in reading rates (see Section 4.6.1 for detailed discussion), but it is beyond the scope of the EPER Placement Test to provide evidence that extensive reading promotes reading rates. Similarly, the participants' progress in terms of EPER level cannot be interpreted as development in automaticity. Automaticity enables the readers to pay attention to content of the text, and build up comprehension by activating background knowledge, and these processes are highly automatized, consuming or occupying little or no attentional resources (Grabe, 2009). In this sense, automaticity is a subcomponent of or contributes to rate, but the EPER Placement Test is not intended to measure the participants' automaticity in reading, either.

As the third component in reading fluency, accuracy "is most closely associated with word recognition in that fluent word recognition must not only be rapid and automatic, it must also be complete and accurate" (Grabe, 2009, p. 291). It differs from the "accuracy" discussed in the previous section, which means the experimental participants comprehend the texts more correctly than the participants in the control group. The "accuracy" for reading according to Kuhn and Stahl (2003) means "confirming what is already known to the reader" (p. 4), agreed by Grabe (2009) that "Fluent readers can recognize almost every word they encounter in a text automatically" (p. 289). In this sense, the participants' improvement in EPER level cannot be interpreted as developing in accuracy, because accuracy means making use of what is already known, while the participants' progress in EPER level means they are able to read materials of increasing levels of difficulty, which involves new words, more complex syntactic structures and more information density.

A fourth component of reading fluency which is increasingly cited is the recognition of prosodic phrasing and contours of the text while reading (e.g.,

Schwanenflugel & Ruston, 2008; Kuhn et al., 2006), which means a fluent reader should be able to pause at appropriate junctures or processing text chunks in ways that match structure units in the continuous prose. As shown in the previous section, the two experimental groups have demonstrated significantly better ability to manage the discourse of the text, namely the flow of the message, but still, it is beyond the scope of the EPER Placement Test.

The mismatch between the progress in EPER level and the core components of reading fluency, however, does not mean that the results of the EPER Placement Test cannot be interpreted within the framework of reading fluency, as the scope and definition of reading fluency has been extended by Nation (2009) that “fluency is....accompanied by improvements in accuracy and complexity” (p. 65). The participants’ progress in EPER level can be interpreted as their development in reading fluency in terms of complexity.

To conclude our discussion in this section, the last area of inquiry to be addressed is the relationship between EPER level and reading comprehension. The two experimental groups significantly outperformed the Control group on both reading comprehension and EPER level. Although the two constructs were measured by two different sets of tests, one developed by the test bureau of a Chinese school district and one developed by an European research institution, they are parameters of the participants’ reading ability. Their reading comprehension performance measured by the pre-, mid and post-tests shows their ability to comprehend short texts, and their EPER levels measured by the EPER test demonstrates their ability to handle long texts. In the present study, both the two experimental groups showed superior ability to handle short and long texts than the Control group. It means there is a positive correlation between reading comprehension and EPER level. Given that the participants’ reading proficiency is improved after this program, they are assumed to manage both short and long texts equally well. On the other hand, there is difference between reading comprehension and EPER level. In the cloze section in the post-test, the Free Reading group was slightly better than the Complementary Reading group in terms of mean score (CR=6.55, FR=6.79), although the difference was not statistically significant ($p=0.51$). The situation was manipulated in the reading

comprehension section (CR=11.55, FR=11.58, $p=0.95$). These two cases show that the Free Reading group performed slightly better on comprehending short texts than the Complementary Reading group judging by mean score, although the difference was not statistically significant. To perform well in these two sections, the participants needed to be accurate readers, by scanning for specific information, by predicting the information flow, and by recognizing the writer's stance. In contrast, the participants needed to be fluent readers to perform well in the EPER test. They should be able to make sense of reading materials of increasing levels of difficulty. The comparison between the two constructs, i.e., reading comprehension and EPER level, does not mean priority should be given to any of them. On the contrary, accuracy and fluency interact with each other in reading process, regardless of the reading purposes. To be a good reader, one should develop both accuracy and fluency in reading. This is what the present study shows to us, and the significance of administering two sets of reading tests to the participants in the present study.

4.5.3 Extensive reading and listening proficiency

As noted and discussed in Chapter 2, there have been a comprehensive range of research on the significance of extensive reading to ESL and EFL teaching and learning showing that extensive reading can effectively improve students' proficiency in various proficiency areas and language knowledge components. Again we do a brief review on the effect of extensive reading as we did in Chapter 2 to refresh our memory. According to Richards et al. (1992), extensive reading is "intended to develop good reading habits, to build up knowledge of vocabulary and structure, and to encourage a liking for reading" (p.133). Grabe and Stoller (2011) propose that benefits from extensive reading include "vocabulary recycling, sight word reading gains, reading speed gains, reading practice, and a habit of being a successful reader in the L2" (p. 122), and argues for extensive reading as a central rather than peripheral component in L2 reading instruction. Besides its role in language proficiency development and reading instruction, "extensive reading plays a key role in moving learners beyond the basic levels of proficiency required for basic communication so that they are well positioned to achieve personal and professional goals" (Horst, 2009, p.63). As noted from the existing body of research, extensive reading seems to be

all-embracing but the relationship between extensive reading and listening proficiency is less discussed. Grabe and Stoller (2011) propose extensive reading improves comprehension skills, without specifying whether extensive reading promotes reading comprehension skills or listening comprehension skills or both. Thus the primary focus of this section is to examine the relationship between extensive reading and listening proficiency based on the test results, with an aim to provide insight into this less-discussed area.

Discussion 1: Are extensive reading and listening proficiency related?

Before starting Discussion One, we must determine that two conditions are equal for the three groups in the present study. The first condition should be that the three groups entered the present study with approximately equal listening proficiency. However, there was not a listening component in the pre-test, and thus the pre-test was unable to provide us with specific or exact information on the participants' listening proficiency at the beginning of this program. The listening proficiency demonstrated by the three groups at the mid of the present study was still roughly equal. As measured by the mid-test, there was no significant difference in listening among the three groups ($F=0.28$, $p=0.76$) (see Table 4.45 below). Thus regarding the first condition, we do not have indicator of the three groups' beginning listening proficiency, as the pre-test provided no assessment of this. The premise of our discussion is that the three groups' listening proficiency was approximately equal from the mid-test.

The second condition should be that the three groups received approximately equal amount of written and oral input; otherwise, their improvement in listening proficiency will be unable to be attributed to extensive reading. In Chapter 3, it has been shown that the EFL context and the homogeneity of the sample make for a great deal of predictability in terms of participants' exposure to English. In the case of foreign language learning, the classroom may be the only source of reliable audio input in the target language for learners at all levels. As discussed in Chapter 1, the Grammar-Translation Method dominates in-class English language teaching in mainland China. Using this method, English is taught in students' mother tongue, i.e., Chinese. The use of Chinese as the dominant medium of instruction effectively avoids the problems in CBEI as introduced in

Chapter 1, i.e., the students' Chinese development will not suffer and their content area knowledge do not need to be simplified or reduced. The third benefit of teaching English in Chinese is that even teachers who do not have a level of English language proficiency can teach English using this method, which is particularly appealing when there are not enough qualified English teachers.

In this situation, the Grammar-Translation Method limited the amount of audio input that the Free Reading group and the Control group can obtain from the classroom. The situation also applied to the Complementary Reading group. Although the Complementary Reading group discussed grammar and translation less in class, their discussion on extensive reading was done in Chinese, their mother tongue. They did not receive extra audio input, either. Thus it is ensured that the participants received nearly the same amount of audio input in the program. The change in their listening proficiency was attributed to extensive reading.

As shown in Table 4.45, there was no significant main effect for the treatments as a factor in the listening section in the mid-test ($F=0.28$, $p=0.76$), i.e., in the middle of the program, but there was a significant main effect for the treatments as a factor in the listening section in the post-test ($F=3.38$, $p=0.04$).

Table 4. 45 Descriptive statistics and results of ANOVA from the listening sections in the mid- and post-tests

| Tests (Total score) | Group | N | Mean | F | Sig. |
|---------------------|-------|----|-------|------|------|
| Mid-test (20) | CR | 33 | 15.24 | 0.28 | 0.76 |
| | FR | 33 | 15.42 | | |
| | CG | 33 | 15.00 | | |
| Post-test (20) | CR | 33 | 14.48 | 3.38 | 0.04 |
| | FR | 33 | 15.42 | | |
| | CG | 33 | 14.06 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Judging from the mean scores in the post-test, the Free Reading group was the best at listening among the three. As shown in Table 4.46 below, the multiple comparisons by the Fisher's LSD test illustrated that the listening score of the Free Reading group was statistically significantly higher than that of the Control Group ($FR=15.42$, $CG=14.06$, $p=0.01$). There was no statistically significant

difference in listening between the Complementary Reading group and the Control Group (CR=14.48, CG=14.06, $p=0.43$), or the two experimental groups (CR=14.48, FR=15.42, $p=0.08$)

Table 4. 46 The results of multiple comparisons by the Fisher's LSD test in the listening section in the post-test

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| Listening | CR | FR | -0.94 | 0.54 | 0.08 |
| | | CG | 0.42 | 0.54 | 0.43 |
| | FR | CR | 0.94 | 0.54 | 0.08 |
| | | CG | 1.36 | 0.54 | 0.01 |
| | CG | CR | -0.42 | 0.54 | 0.43 |
| | | FR | -1.36 | 0.54 | 0.01 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

The post-test result appeared to provide solid evidence that there was a positive relationship between extensive reading and listening proficiency. However, it is too early to confirm that extensive reading and listening are related solely based on this study. The research literature concerning extensive reading produced conflicting results supporting either positive or null relationship between extensive reading and listening proficiency.

The study conducted by Lai (1993b), in which 1,351 secondary school students participated in either a four-week summer reading program or a year-long reading scheme, showed that extensive reading resulted in significant gains in listening comprehension. Conversely, in an extensive reading program implemented in Fujian Normal University, which compared the effect of the Extensive Reading Approach and the Intensive Reading Approach, i.e., the Grammar-Translation Method, on improving students' English language proficiency, the findings showed that the two different approaches made no difference in improving students' listening proficiency ($p=.602$). On the contrary, in the Fujian program the Intensive Reading group was slightly better in this proficiency area with the Intensive Reading group 10.20 and the Extensive Reading group 9.80 (Tan, 1998, p. 112). The conflicting results produced by the empirical studies provide an optimal chance for the comparison of the specific conditions of different extensive reading programs to seek the root causes of the positive or null relationship between extensive reading and listening proficiency.

For our discussion, the comparison between the Fujian program and the present study are of particular interest, as they produced conflicting results on the relationship between extensive reading and listening proficiency. In the Fujian program, the participants read around one million words, and in the present study, the students in the Complementary Reading group read around one million words as well. Then, what are the reasons leading to different results in two extensive reading programs with a ten-year interval?

In the Fujian program, the four hours of college English teaching time per week was divided into two halves: two hours for normal teaching and two hours for the literature circle, which was defined as self-selected discussion groups that centered on commonly read stories, articles and books. Four to five students reading the same book were in one discussion group. Every group recommended a group leader, who was responsible for reporting to the teacher about what they had chosen to read and organizing the discussion. Since the students were all very poor in spoken English, the discussion was allowed to go on in Chinese so that the students could participate in the discussion without being embarrassed by their poor spoken English. Through the group discussion, the students could have a better understanding of the books by hearing different voices and critiques and by sharing their experience in solving the problems they met with in their reading. In the present study, the teaching time was five sessions of forty minutes per week, namely 200 minutes, 3.2 hours. For the Free Reading group, all the 200 minutes was devoted to normal teaching using the Grammar-Translation Method. For the Complementary Reading group, the time devoted to in-class discussion accounted for only 10% to 15% of the total teaching time.

In comparison, the Fujian program shared something in common with the Complementary Reading group in the present study: a large reading amount of more than one million, out-of-class extensive reading and in-class discussion in Chinese. The similarity produced similar results that extensive reading did not result in significantly better results in reading. In contrast, the Free Reading treatment, which was defined as an extracurricular reading activity with no time devoted to in-class discussion, led to significant improvement in listening.

As shown in Table 4.47 below, there are three key differences between the Fujian program and the Free Reading treatment in the present study. The first key difference was that the reading materials in the Fujian program was unsimplified novels, magazines and newspapers whereas the reading material in the latter was graded readers. The second key difference was that the free reading participants in the present study did not discuss their reading in groups; instead, extensive reading was an individualized activity. The third key difference was that the participants in the Fujian program read approximately eight times as much as that of the free reading participants in the present study.

Thus, putting these conflicting results together, we have reason to believe that there is indeed a positive relationship between extensive reading and listening proficiency, but the key factor cannot be in-class discussion or reading quantity. Our next step is to further discuss how extensive reading and listening are related with an effort to identify contributing factors to listening proficiency in extensive reading.

Discussion 2: How are extensive reading and listening related?

In the Fujian program, extensive reading did not produce better result on listening compared with intensive reading. Tan's (1998) explanation for the result was that listening and reading differ as skills in communication because listening belongs to the category of spoken language and reading the written language. His explanation appeared to be reasonable from this perspective; however, from another perspective reading and listening are categorized as receptive skills. It is hardly convincing that reading and listening are independent of each other.

To identify the factors contributing to the improvement of listening proficiency, we first need to explore how listening works. Different theories have been proposed to explore the listening process, among which Anderson's (1995) and Buck's (2001) are particularly influential. Anderson (1995) decomposes listening into a three-stage cognitive processing. In this first stage, *perceptual processing*, phonemes are segmented from speech and these segmented phonemes are stored in listeners' "echoic memory" (p. 137). In the second stage, *parsing*, the

Table 4. 47 Comparison between the Fujian Program and the present study

| Details | Fujian Program | Shanghai Program | |
|---|---|--|--|
| | | CR | FR |
| Level | College | Senior high school | Senior high school |
| Duration | One academic year | One academic year | One academic year |
| Number of participants | 20 | 33 | 33 |
| Reading materials | Novels; magazines and newspapers published in China | OBW graded readers | OBW graded readers |
| Average reading amount | 1,600,000 words | 1,144,279 words | 226,732 words |
| Time for teaching per week | 4 hours | 5 sessions of 40 minutes | 5 sessions of 40 minutes |
| Time for in-class reading | No. | No. | No. |
| Time for in-class discussion | 2 hours per week | 20 to 30 minutes per week | No. |
| Medium of discussion | Chinese | Chinese | / |
| Instrument | College English Test, Band 4 | Tests developed by the school district | Tests developed by the school district |
| Significant better results in listening | No. | No. | Yes. |

segmented phonemes stored in listeners' echoic memory are matched with the knowledge foundation in listeners' long-term memory. In the final stage, *utilization*, the segmented phonemes stored in echoic memory and the knowledge foundation in long-term memory work in listeners' schemata to form meaning. Unlike Anderson's linear model, Buck (2001) describes listening as an interaction between two knowledge resources, namely, the listeners' background knowledge and the listeners' linguistic knowledge. These two sets of resources interact to make sense of the listening text. The general consensus between the two theories is that three sets of resources: the listening text, the linguistic knowledge and the schematic knowledge, are needed for processing a listening task.

In keeping with the emphasis on linguistic and schematic knowledge, much of the relevant research numerates other factors that affect listening, of which Rubin (1994) identifies the following variables:

1. the text
2. the task
3. the listener
4. the processing

Among the five variables, the first two were pre-set by the test developers. In the present study, the mid- and post-tests were developed by the school district test bureau. The researcher had no input into or influence over the listening texts employed in the tests. In the tests, the participants listened to the audio recording and answered the questions as required. There was no interlocutor involved in the whole process. The listening sections in the mid- and post-tests were in strict accordance with the format of the listening section in the NMET. In both the mid- and post-tests, listening section was divided into three subsections: listening to the short conversations, listening to the short passages and listening to the longer conversations (see Appendix B and C). The test takers were required to strictly follow the recording and complete the pre-set tasks. Given that the text characteristics, interlocutor characteristics and task characteristics were beyond the scope and control of the present study, our discussion will focus on the latter two characteristics, namely listener characteristic and process characteristic in

subsequent discussion.

There are a considerable number of variables involved the listener characteristics: language proficiency level, learning disabilities, general knowledge, age, gender, motivation, confidence, etc (Rubin, 1994; Vandergrift, 2005). We just note a few here because of limited space.

In Vandergrift's (2006) study, both L1 listening comprehension ability and L2 proficiency are important variables that may affect L2 listening comprehension ability, with L2 proficiency play greater role in L2 listening comprehension as measured by the multiple regression analysis. In the present study, the test results showed that extensive reading improved the participants' overall English language proficiency. We speculate that the improvement in global English language proficiency in turn contributed to listening comprehension ability in English. This speculation is supported by the detailed analysis of the three groups' performance in the listening section in the post-test as presented in Table 4.48 below. As shown in Table 4.48, there was no statistically significant difference among the three groups on individual items except Item 12. In terms of mean score, the three groups showed advantage on different items: the Complementary Reading group performed the best on Items 5, 9, 13, 14, 15, 18 and 19, the Free Reading group took lead on Items 1, 3, 4, 6, 7, 8, 9, 10, 11, 12, 16 and 17, and the Control group outperformed the two experimental groups on Items 2 and 18. This distribution of the three groups' performance on individual items shows that none of the three groups obtained absolute advantage on every item in the listening section in the post-test, and thus the statistically significant difference among the three groups in the whole section can be interpreted as the two experimental groups' advantage in the listening section as a whole or their improvement in overall listening proficiency because of extensive reading. This is a conclusion that we may make at this stage, but we are unable to identify the proportion that improvement of overall L2 proficiency accounts for in the improvement in overall listening proficiency. Similarly, we are unable to answer the question: whether the improvement of listening proficiency is the major effect or a by-product of extensive reading.

Table 4. 48 Descriptive statistics and results of ANOVA from the listening section in the post-test

| Item | Group | N | Mean | SD | F | Sig. |
|------|-------|----|------|------|------|------|
| 1 | CR | 33 | 0.70 | 0.47 | 1.26 | 0.29 |
| | FR | 33 | 0.85 | 0.36 | | |
| | CG | 33 | 0.82 | 0.39 | | |
| 2 | CR | 33 | 0.52 | 0.51 | 0.52 | 0.59 |
| | FR | 33 | 0.55 | 0.51 | | |
| | CG | 33 | 0.64 | 0.49 | | |
| 3 | CR | 33 | 0.27 | 0.45 | 2.23 | 0.11 |
| | FR | 33 | 0.48 | 0.51 | | |
| | CG | 33 | 0.27 | 0.45 | | |
| 4 | CR | 33 | 0.70 | 0.47 | 2.34 | 0.10 |
| | FR | 33 | 0.88 | 0.33 | | |
| | CG | 33 | 0.67 | 0.48 | | |
| 5 | CR | 33 | 0.61 | 0.50 | 2.68 | 0.07 |
| | FR | 33 | 0.36 | 0.49 | | |
| | CG | 33 | 0.36 | 0.49 | | |
| 6 | CR | 33 | 0.82 | 0.39 | 2.01 | 0.14 |
| | FR | 33 | 0.97 | 0.17 | | |
| | CG | 33 | 0.85 | 0.36 | | |
| 7 | CR | 33 | 0.06 | 0.24 | 1.13 | 0.33 |
| | FR | 33 | 0.18 | 0.39 | | |
| | CG | 33 | 0.12 | 0.33 | | |
| 8 | CR | 33 | 0.79 | 0.42 | 0.49 | 0.62 |
| | FR | 33 | 0.88 | 0.33 | | |
| | CG | 33 | 0.82 | 0.39 | | |
| 9 | CR | 33 | 0.94 | 0.24 | 0.21 | 0.81 |
| | FR | 33 | 0.97 | 0.17 | | |
| | CG | 33 | 0.94 | 0.24 | | |
| 10 | CR | 33 | 0.91 | 0.29 | 0.13 | 0.88 |
| | FR | 33 | 0.94 | 0.24 | | |
| | CG | 33 | 0.91 | 0.29 | | |
| 11 | CR | 33 | 0.91 | 0.29 | 0.77 | 0.47 |
| | FR | 33 | 0.94 | 0.24 | | |
| | CG | 33 | 0.85 | 0.36 | | |
| 12 | CR | 33 | 0.48 | 0.51 | 3.32 | 0.04 |
| | FR | 33 | 0.70 | 0.47 | | |
| | CG | 33 | 0.39 | 0.50 | | |
| 13 | CR | 33 | 0.85 | 0.36 | 1.03 | 0.36 |
| | FR | 33 | 0.85 | 0.36 | | |
| | CG | 33 | 0.73 | 0.45 | | |
| 14 | CR | 33 | 0.42 | 0.50 | 0.51 | 0.60 |
| | FR | 33 | 0.36 | 0.49 | | |
| | CG | 33 | 0.30 | 0.47 | | |
| 15 | CR | 33 | 0.82 | 0.39 | 1.46 | 0.24 |
| | FR | 33 | 0.76 | 0.44 | | |
| | CG | 33 | 0.64 | 0.49 | | |
| 16 | CR | 33 | 0.88 | 0.33 | 2.1 | 0.12 |
| | FR | 33 | 1.00 | 0.00 | | |
| | CG | 33 | 0.94 | 0.24 | | |
| 17 | CR | 33 | 0.91 | 0.29 | 1.60 | 0.21 |
| | FR | 33 | 1.00 | 0.00 | | |
| | CG | 33 | 0.91 | 0.29 | | |
| 18 | CR | 33 | 0.94 | 0.24 | 0.53 | 0.59 |
| | FR | 33 | 0.88 | 0.33 | | |
| | CG | 33 | 0.94 | 0.24 | | |
| 19 | CR | 33 | 1.00 | 0.00 | 1.02 | 0.36 |
| | FR | 33 | 0.97 | 0.17 | | |
| | CG | 33 | 0.94 | 0.24 | | |
| 20 | CR | 33 | 0.97 | 0.17 | 0.00 | 1.00 |
| | FR | 33 | 0.97 | 0.17 | | |
| | CG | 33 | 0.97 | 0.17 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

In addition to overall language proficiency, memory is another factor that may affect listening proficiency. The relationship between listening comprehension and memory has been verified by Florit et al. (2009). Memory is usually divided into long-term memory and working memory (Kurt, 2008). While long-term memory is the major source of the linguistic and background information to support reading and listening, the key concept for comprehension is working memory.

A brief explanation of how working memory contributes to reading comprehension may deepen our understanding of its role in listening comprehension (see Baddeley, 2006, 2007; Cain, 2006; Long, et al., 2006, for detailed discussion of the relationship between working memory and reading comprehension; see also Kintsch, 1988; Gernsbacher, 1997; Just & Carpenter, 1992, for the role of working memory in specific models of reading). As introduced in Section 1.1, reading comprehension activates a series of processes: (1) word recognition (recognition of words in the text), (2) syntactic parsing (extracting grammatical information from a group of words, which is crucial for disambiguating polysemy in word recognition), and (3) semantic proposition formation, i.e., building clause-level meaning on the basis of word recognition and syntactic parsing. This set of processes is commonly referred to as bottom-up processing (Alderson, 2000) or lower-level processing (Grabe, 2009; Grabe & Stoller, 2011). The lower-level processes are language-oriented, which require the readers' knowledge about the language in which they read. Based on the knowledge-oriented lower-level processes, another set of processes that contributes to comprehension is activated. After the clause-level meanings are built, they are integrated into a broader network in which the clause-level meanings are connected and the main idea of the text is formed. This process is referred to as Text Model of Comprehension, a summary of the text itself. Spontaneously, readers will interpret the text in terms of their reading purposes and background knowledge, etc. This process is called the Situational Model of Reader Interpretation, in which a number of reader variables noted above, such

as knowledge of the world and cultural knowledge, have an important role to play. The description of how the lower-level and higher-level processes are carried out does not imply that they are carried out independently at different times. It is in working memory that these two sets of processes are integrated and comprehension occurs.

In this process, working memory is not a concrete thing like a “box” located somewhere in the brain. It is proposed that working memory is best understood as a hypothesis that explains how comprehension occurs in our brain (e.g., Baddeley, 2007). An analogy can be drawn on between working memory and platform that working memory is a “platform” on which the brain processes information and builds comprehension. This hypothesis points out the two core functions of working memory: storage, i.e., the “platform” should be “wide” enough to accommodate the information pending processing or being processed, and processing, i.e., how the brain works with the information on this “platform”.

Following this logic, improved working memory by extensive reading may contribute to listening comprehension in two possible ways: greater storage capacity to keep the information active for a longer period of time, or better processing function to process the information more effectively. The first possibility is consistent with Anderson’s (1995) theory of listening as described earlier in this section. In Anderson’s theory, in the first stage of listening, *perceptual processing*, listeners segment phonemes from the speech and these segmented phonemes are stored in “echoic memory” (p. 137). The “echoic memory” is exactly the “storage” component of working memory.

The detailed analysis of the three groups’ performance on individual items in the listening section in the post-test presented in Table 4.49 may support the first possibility, i.e., greater storage capacity in working memory after extensive reading. As shown in Table 4.48 above, there is a significant main effect for the treatments as a factor on Item 12 in Section B Passages. In Section B, a passage will first be read twice and then three questions based on the passage will be read once. That is to say, the information is prior to questions. In this situation, the test takers may not be able to predict what questions will be asked, and thus they will

not be able to locate information for a specific issue. Preferably, the more information they can restore, the better they can perform better in this section. If the storage capacity in the working memory of the test takers is not great enough to keep what have been read active for sufficient duration, they will be unable to remember or process what have been read when the question is raised. The post hoc multiple comparisons by the Fisher's LSD test detected that the statistically significant difference on Item 12 is between the Free Reading group and the Control Group (FR=0.70, CG=0.39, $p=0.01$) (see Table 4.49 below) and the Complementary Reading group performed better on Item 12 than the Control group in terms of mean score (CR=0.48, CG=0.39, $p=0.45$), which suggests that the participants in the Complementary Reading group may greater storage capacity in their working memory than the participants in the Control group.

Table 4. 49 The results of multiple comparisons by the Fisher's LSD test on Item 12 in the listening section of the post-test

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| Item 12 | CR | FR | -0.21 | 0.12 | 0.08 |
| | | CG | 0.09 | 0.12 | 0.45 |
| | FR | CR | 0.21 | 0.12 | 0.08 |
| | | CG | 0.30 | 0.12 | 0.01 |
| | CG | CR | -0.09 | 0.12 | 0.45 |
| | | FR | -0.30 | 0.12 | 0.01 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

The significance of storage capacity in working memory is particularly important for listening comprehension. In reading comprehension, if the readers are not able to store too much information in their working memory for processing, they will still have the chance of rereading because the reading materials are available to them all the time. In comparison, listeners will not have the change of re-listening because of the complexity and rapidity of the listening process. In this situation, a greater storage capacity in listeners' working memory will greatly facilitate listening comprehension.

One of the limitations of the present study in exploring the relationship between extensive reading and listening comprehension is that it does not find the answer to this question: How the participants' storage capacity in working memory were expanded after extensive reading? At this stage, a preliminary explanation is that

the participants read lengthy instead of short texts in the program, and thus they would need to keep information active in their working memory; otherwise they would be able to building meaning of the whole story without integrating previous and updated information. How this improvement happened is not visible to naked eyes, as Alderson and Urquhart (1984) point out, “Research into reading is research into the unobservable” (p. x) (see Section 2.1.1 for this statement). The second limitation of the present study in exploring the relationship between extensive reading and listening comprehension is that the greater storage capacity in working memory after extensive reading is one of the possibilities that may explain the participants’ performance in the listening section in the post-test. It cannot exclude the possibility that the other component in working memory, *processing*, contributed to the participants’ listening performance; similarly, the proportion of contributed made by storage capacity and processing cannot be worked out or quantified.

Discussion 3: What else contributed to the improvement of listening proficiency?

A test-taker’s performance is affected by at least two variables: the test taker’s ability and the test technique (Wu, 1998). Test techniques are particularly important in assessing listening. Because of the complexity and rapidity of listening process, the cognitive processes involved in interpreting a listening text are invisible. The test takers’ listening performance is largely decided by the type of responses required.

It has been confirmed that test formats have a significant effect on test takers’ performance (e.g., In’nami & Koizumi, 2009). In Cheng’s (2004) study, there test formats, i.e., traditional multiple-choice, multiple-choice cloze and open-ended questions, were employed to assessing the participants’ listening comprehension. The 159 Chinese EFL university students in this study performed best on the multiple-choice cloze, and scored lowest on the open-ended questions. In the post-study survey, the overwhelming majority (97%) of the participants pointed out that the multiple-choice formats were preferred to open-end questions, because the options given in the multiple-choice formats aided comprehension and improved accuracy in guessing.

In fact, the Chinese EFL learners' preference for the multiple choice format is not good news for every test taker. The multiple choice format threatens the construct validity of listening assessment in the following two ways. First, it favors the more advanced listener, but disadvantages the less advanced. Second, it connives uninformed guessing and allows the test takers to give the correct answer for the wrong reasons (Wu, 1998). Wu's finding helped to partially explain why there was no significant difference in the present study among the three groups in listening in the mid-test but in the post-test the significant difference came as a surprise. As evidenced by the mid-test, there was no statistically significant difference among the two experimental groups and the Control group in five out of seven comparisons in the mid-test, which indicated that the two experimental groups had not obtained absolute advantage by the mid-test. By the end of the program, the two experimental groups outperformed the Control group on five out of the seven comparisons in the post-test, which meant the two experimental groups, especially the Free Reading group, have obtained advantage in overall language proficiency, and specifically, in listening. The two experimental groups, or at least the Free Reading group, were more likely to be at an advantage in the listening section in the post-test according to Wu's study.

If so, the multiple-choice fact should not be the dominant test format in high-stakes English tests. In fact, the multiple choice format is not the only suggested test format for listening. As specified in the 2001 English Curriculum Standards (MOE, 2001), several possible test formats are suggested, as shown in Table 4.50 below.

Table 4. 50 Suggested test points and formats for listening comprehension in the 2001 syllabus

| Test points | | Test formats | | Explanation |
|-------------|---|--------------|--|---|
| 1. | Obtain specific and factual information. | 1. | Multiple choices. | The language materials selected must reflect the characteristics of true oral communication. They should be medium in length and be read at normal speech rate. |
| 2. | Understand the subject and the gist. | 2. | True or false. | |
| 3. | Make inference about the background of the content and the relationship between the speakers. | 3. | Complete the sentence. | |
| | | 4. | Complete the chart and diagram. | |
| 4. | Understand the speaker's intentions, views or attitudes. | 5. | Make simple record and summary. | |
| | | 6. | Complete tasks according to the instruction. | |

Despite the several suggested test formats, the multiple-choice format is the most widely adopted test format in both local and national tests. It will co-exist with subsequent extensive reading research in this context for long in the future. Subsequent extensive reading research in this context may take this factor into consideration while selecting or developing tests for extensive reading programs.

Discussion 4: What is the significance of extensive reading to listening proficiency?

The relationship between extensive reading and listening proficiency has been investigated and discussed in the previous three sections. Our focus on listening does not only lie in its role in our study but also its importance in second language acquisition (e.g., Elley, 1989). The importance of input for second language acquisition has been stressed by second language acquisition theory. Linguistic input can be either in a written or oral form and thus listening and reading is of equal importance in receiving linguistic input. Furthermore, listening and reading is complementary to each other as they are categorized as receptive skills. It is suggested by Long (1986) that students will gain greater confidence when both of their receptive skills are well developed. Similarly, Rivers (1981) proposed that students should be able to comprehend all types of input from the early stages of language learning. Ultimately, the enhancement of students' ability to hear and read in a second or foreign language will facilitate their development of second or foreign language comprehension (e.g., Mills, et al., 2006; Park, 2004).

However, the importance of listening in mainland secondary school has not been given due attention despite the significance of the enhancement of students' listening proficiency as presented in the previous paragraph. In a Grammar-Translation Method dominant classroom, the time that can be devoted to listening is understandably low. The strongest motivation for the teachers to improve students' listening proficiency comes from the washback effect of the NMET, the highest-takes English test in mainland China. In the NMET, listening accounts for 20% of the total score, which means students' improvement in listening will certainly lead to the improvement of total scores.

Consequently, listening seems to deserve a place in classroom teaching, and there has been an argument whether listening is teachable. One group of researchers suspects that listening cannot be taught. The other group of researchers insists that teaching students how to listen does make a difference (e.g., Vandergrift & Tafaghodtari, 2010). For example, Thompson and Rubin (1996) confirm that systematic instruction of listening strategy lead to the improvement of listening comprehension.

What the present study contributes to this argument is that it provides empirical evidence supporting the positive relationship between extensive reading and listening proficiency, which suggests one more possible way for the improvement of listening proficiency. Of particular value is that extensive reading is mostly done out of class, which is an optimal supplement to limited instructional time as presented in Chapter 1.

4.6 Extensive reading and productive skills

4.6.1 Extensive reading and translation

Strictly speaking, translation is not a language skill. The four recognized language skills are listening, reading, speaking and writing. The first two are referred to as receptive skills and the latter two productive skills. The justification for defining translation as a productive skill is that the students are asked to produce grammatically correct sentences as required, without four options provided for them to choose from. In this sense, translation in the present study can be considered as a writing task, although distorted somehow.

Although translation is a writing task, there is no room for the students to show their personal thinking, their size of vocabulary, and their lexical choices nor are they able to display any sense of wit and humor. In the translation section in the post-test, the students were required to translate five Chinese sentences into English, with three points for each sentence. One English word was given for each sentence in the bracket without specifying its part of speech. The test takers would lose one point if they did not use the word given in the bracket in their translation. Grammatical and spelling mistakes also led to loss of marks (see Appendices A, B and C).

To construct a correct sentence, the participants need to avoid any grammar or lexical mistakes, which will lead to loss of marks. For instance, the missing of “s” from the end of the verb after the third person singular, “He like swimming”, will lead to a loss of one point because the absence of “s” is considered to be a grammar mistake. If another student writes “He likes swiming”, half a point will be deducted, because the absence of “m” is considered to be a spelling mistake. That a grammar mistake costs one point and a spelling mistake costs half a point is a commonly shared rule among the high school English teachers in Shanghai, and this was exactly how the translation parts in the pre-, mid- and post-tests were marked (see Section 3.3 for the marking procedures).

In fact the two sentences “He like swimming” and “He likes swiming” convey the same meaning despite the minor mistakes, but when marked they were indeed treated differently. Given that greater attention was given to accuracy of form in the translation task, the English teachers in Shanghai secondary schools pay special attention to translation exercises because a minor mistake will lead to loss of marks. Repeated exercises will be needed if a student wants to achieve a satisfactory score in the translation section, and translation exercises are usually an integral part of the students’ daily assignment.

Table 4.51 below clearly shows how the Complementary Reading group lost their advantage in this section gradually. In the pre-test, the Complementary Reading group was ahead of the Free Reading group by 0.4 point and the Control group by 0.5 point when the full mark of this section was ten points. In the mid-test, the gap between the Complementary Reading group and the Free Reading group was narrowed to less than 0.3 points when the full mark of this section was 15 points. At that stage, an excuse could be found as the Complementary Reading group still kept its advantage over the Control group by 0.8 points. However, the situation in the post-test was a 180-degree turn to that in the pre-test. The Free Reading group and the Control group outperformed the Complementary Reading group, and there is a significant main effect for the treatments as a factor in this section ($F=6.48, p<0.001$).

Table 4. 51 Descriptive statistics and results of ANOVA from the translation sections in the pre-, mid- and post-tests

| Tests (Total score) | Group | N | Mean | SD | F | Sig. |
|---------------------|-------|----|-------|------|------|------|
| Pre-test (10) | CR | 33 | 7.03 | 1.59 | 0.88 | 0.42 |
| | FR | 33 | 6.64 | 1.65 | | |
| | CG | 33 | 6.52 | 1.70 | | |
| Mid-test (15) | CR | 33 | 11.92 | 1.17 | 1.53 | 0.22 |
| | FR | 33 | 11.68 | 1.62 | | |
| | CG | 33 | 11.15 | 2.47 | | |
| Post-test (15) | CR | 33 | 10.55 | 1.42 | 6.48 | 0.00 |
| | FR | 33 | 11.76 | 1.40 | | |
| | CG | 33 | 11.15 | 1.28 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

Table 4.52 below confirms that the Complementary Reading group was in an inferior position in the translation section in the post-test. The Free Reading group outperformed the Complementary Reading group by 1.21 points (CR=10.55, FR=11.76), and this difference between the two groups was statistically significant ($p<0.001$). Although the difference between the Complementary Reading group and the Control group was not statistically significant ($p=0.07$), the Control group had an advantage of 0.61 point. Given that the Complementary Reading group entered this program with better translation performance, this advantage could be bigger than 0.61.

Table 4. 52 The results of multiple comparisons by the Fisher's LSD test in the translation sections in the pre-, mid- and post-tests

| Tests (Total score) | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|---------------------|-----------|-----------|-----------------------|------------|------|
| Pre-test (10) | CR | FR | 0.39 | 0.41 | 0.34 |
| | | CG | 0.42 | 0.41 | 0.31 |
| | FR | CR | -0.39 | 0.41 | 0.34 |
| | | CG | 0.03 | 0.41 | 0.94 |
| | CG | CR | -0.42 | 0.41 | 0.31 |
| | | FR | -0.03 | 0.41 | 0.94 |
| Mid-test (15) | CR | FR | 0.24 | 0.45 | 0.59 |
| | | CG | 0.77 | 0.45 | 0.09 |
| | FR | CR | -0.24 | 0.45 | 0.59 |
| | | CG | 0.53 | 0.45 | 0.24 |
| | CG | CR | -0.77 | 0.45 | 0.09 |
| | | FR | -0.53 | 0.45 | 0.24 |
| Post-test (15) | CR | FR | -1.21 | 0.34 | 0.00 |
| | | CG | -0.61 | 0.34 | 0.07 |
| | FR | CR | 1.21 | 0.34 | 0.00 |
| | | CG | 0.61 | 0.34 | 0.07 |
| | CG | CR | 0.61 | 0.34 | 0.07 |
| | | FR | -0.61 | 0.34 | 0.07 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

The inferior position of the Complementary Reading group, however, did not announce the ineffectiveness or failure of extensive reading in enhancing productive skills. The Free Reading group performed the best among the three groups, and was ahead of the Complementary Reading group by 1.21 points and the Control group by 0.61 point. The success of the Free Reading group demonstrates that extensive reading is not useless even in a highly accuracy-focused teaching and testing environment. At the same time, the findings also indicate the importance of attention to accuracy if students are to improve in translation.

The detailed examination of the three groups' performance on the five sentences in the translation section in the post-test presented in Table 4.53 below shows that the only statistical significance in the translation section was on Sentence 2, whose correct answer was: *There is a lot in common between the Chinese and Japanese cultures/China and Japan share a lot in common in their cultures*. The key word given for this translation is *common*, without specifying its part of speech. *Common* appeared in Chapter 4 in the second academic term. It was explicitly taught as a core word in class. All the students were requested to learn the phrase by heart: *have... in common*, and translation excises were designed to consolidate students' understanding of this word.

In fact, the five key words tested in the translation section in the post-test, i.e., *success*, *common*, *so...that*, *guilty* and *emphasize* appeared in the textbook and were explicitly taught in class. Thus one reason which may explain why the Complementary Reading group was the weakest in the translation section was that the complementary reading participants seldom did written assignments during the present study because they kept reading every day. This lack of practice understandably led to their inferior position to the other two groups.

A further detailed analysis of the translation section in the post-test presented in Table 4.54 shows that there was no statistically significant difference between each group on Sentence 1, 3, 4 and 5, although the Complementary Reading group lagged behind if judged by net scores. As introduced in this section earlier, both grammatical and spelling mistakes in the sentences would lead to loss of

marks. The multiple comparisons presented in Table 4.52 have shown that there was no statistically significant difference between the three groups on most sentences. It is reasonable to assume that the Complementary Reading group was able to construct meaning-correct sentences as the other two groups. It was highly probable that it was their mistakes in accuracy that resulted in their poor performance.

Table 4. 53 Descriptive statistics and results of ANOVA from the translation section in the post-test

| Item | Group | N | Mean | SD | F | Sig. |
|------|-------|----|------|------|------|------|
| 1 | CR | 33 | 2.76 | 0.38 | 1.47 | 0.24 |
| | FR | 33 | 2.89 | 0.24 | | |
| | CG | 33 | 2.80 | 0.35 | | |
| 2 | CR | 33 | 1.67 | 0.97 | 5.48 | 0.01 |
| | FR | 33 | 2.29 | 0.59 | | |
| | CG | 33 | 2.03 | 0.68 | | |
| 3 | CR | 33 | 2.42 | 0.57 | 1.35 | 0.26 |
| | FR | 33 | 2.53 | 0.43 | | |
| | CG | 33 | 2.32 | 0.56 | | |
| 4 | CR | 33 | 2.14 | 0.50 | 1.49 | 0.23 |
| | FR | 33 | 2.23 | 0.61 | | |
| | CG | 33 | 2.36 | 0.49 | | |
| 5 | CR | 33 | 1.56 | 0.56 | 1.50 | 0.23 |
| | FR | 33 | 1.82 | 0.67 | | |
| | CG | 33 | 1.64 | 0.63 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

To summarize, the Complementary Reading group's performance in the translation section in the post-test shows that it is too optimistic of us to hypothesize that the effect of extensive reading is universal. In fact, extensive reading faces a bigger challenge in contexts where the classroom instruction and exams focus on accuracy rather than fluency.

Accuracy and fluency, the two sides of one coin, have often led to disputes in language teaching and learning. Logically, the relationship between the focus on accuracy and focus on fluency has at least the following five possibilities: (1) focus primarily on accuracy; (2) focus on accuracy with affordable attention to fluency; (3) equal focus on accuracy and fluency; (4) focus on fluency with affordable attention to accuracy; (5) focus primarily on fluency. In teaching practice, the five possibilities are constrained by different concepts of language and different concepts of language teaching and learning. Different teaching

purposes, teaching objectives and teaching environment will often lead to different conclusions in different empirical studies.

Table 4. 54 The results of multiple comparisons by the Fisher's LSD test in the translation section in the post-test

| Item | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|------|-----------|-----------|-----------------------|------------|------|
| 1 | CR | FR | -0.14 | 0.08 | 0.10 |
| | | CG | -0.05 | 0.08 | 0.58 |
| | FR | CR | 0.14 | 0.08 | 0.10 |
| | | CG | 0.09 | 0.08 | 0.26 |
| | CG | CR | 0.05 | 0.08 | 0.58 |
| | | FR | -0.09 | 0.08 | 0.26 |
| 2 | CR | FR | -0.62 | 0.19 | 0.00 |
| | | CG | -0.36 | 0.19 | 0.06 |
| | FR | CR | 0.62 | 0.19 | 0.00 |
| | | CG | 0.26 | 0.19 | 0.18 |
| | CG | CR | 0.36 | 0.19 | 0.06 |
| | | FR | -0.26 | 0.19 | 0.18 |
| 3 | CR | FR | -0.11 | 0.13 | 0.41 |
| | | CG | 0.11 | 0.13 | 0.41 |
| | FR | CR | 0.11 | 0.13 | 0.41 |
| | | CG | 0.21 | 0.13 | 0.10 |
| | CG | CR | -0.11 | 0.13 | 0.41 |
| | | FR | -0.21 | 0.13 | 0.10 |
| 4 | CR | FR | -0.09 | 0.13 | 0.49 |
| | | CG | -0.23 | 0.13 | 0.09 |
| | FR | CR | 0.09 | 0.13 | 0.49 |
| | | CG | -0.14 | 0.13 | 0.31 |
| | CG | CR | 0.23 | 0.13 | 0.09 |
| | | FR | 0.14 | 0.13 | 0.31 |
| 5 | CR | FR | -0.26 | 0.15 | 0.09 |
| | | CG | -0.08 | 0.15 | 0.62 |
| | FR | CR | 0.26 | 0.15 | 0.09 |
| | | CG | 0.18 | 0.15 | 0.24 |
| | CG | CR | 0.08 | 0.15 | 0.62 |
| | | FR | -0.18 | 0.15 | 0.24 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

In this study, extensive reading and the Grammar-Translation Method are the two ends on the focus on accuracy and focus on fluency continuum. As the dominant English language teaching and learning method in the mainland, the Grammar-Translation Method has its main theoretical basis that second/foreign language learning in the classroom is a cognitive process; it is essentially within the areas of skill learning. The Grammar-Translation Method has three commendable features: (1) it deepens the learners' understand of the grammatical structures in the target language in various ways, for example, through interpreting the grammatical forms of the target language in L1, and through

comparing L1 with the target language; (2) the exercises can be for understanding, i.e. grammar exercises in a multiple choice format, or for output, i.e. translation exercises; (3) The Grammar-Translation exercises provide opportunities to automatize their grammar skills and improve their language accuracy.

The test results of the grammar and translation sections in the post-test present a mixed picture of the ability of extensive reading to promote accuracy development. Our discussion in these two sections focus on two objectives, namely accuracy and fluency: (1) to highlight the central aspects of the effect of extensive reading on developing fluency, and (2) to show the limitations of extensive reading in this accuracy-oriented English language teaching and learning environment.

In addition to its limitations in building accuracy, extensive reading is constrained by many contextual factors in developing fluency. In particular, teachers are obligated to train the student in terms of accuracy, because accuracy is a crucial or even core criterion in high-stakes exams, such as the NMET (see Section 1.4 for introduction of this highest-stakes English exam in mainland China, and see also Section 3.7 for the washback effect of the NMET on English language teaching and learning in secondary schools). Obviously, accuracy development is on top of the teachers' agenda because of the NMET and the whole educational system in this context.

4.6.2 Extensive reading and writing proficiency

In many ways, reading and writing appear to be distinct skills, and in teaching practice they are mostly and traditionally taught separately. Viewing reading and writing as separate skills or processes is a tradition which dates back to American colonial schools (Nelson & Calfee, 1998). However, over the past two decades a close connection has been proposed between the two (e.g., Pearson & Tierney, 1984). Wittrock (1983) summarizes the relationship between reading and writing by pointing out that “Good reading, like effective writing, involves generative processes that create meaning by building relations between the text and what we know, believe, and experience” (p. 600). Smith (1983) focuses on the

reading-to-write relationship by claiming that it is the only way to explain how anyone learns to write. The value of Smith's argument is that it explains how people gradually learn writing. In teaching practice, it is beyond the capacity of explicit instruction to teach the complexities of register, genre, spelling form, appropriateness, etc. He points to the necessity of learning to write from what is read in a subconscious, incidental and collaborative manner. On the relationship between reading and writing, Krashen (1993) made an even stronger claim that "Reading is the only way, the only way we become good readers, develop a good writing style, an adequate vocabulary, advanced grammar, and the only way we become good spellers" (p. 23). Krashen's assumption can be interpreted as readers will internalize the style, structure and correctness by means of exposure to print (Hudson, 2007).

These arguments for the connection between reading and writing are supported by the existing body of literature. These studies can be divided into two groups: reading-to-write task and reading and writing relationship. In the first research area, the writing tasks are based on reading. Kennedy (1985) studied three fluent and three less fluent college students as they constructed an essay after reading three assigned articles. The essay should be based the material contained in the three articles. The first finding was that the six participants finished the writing tasks in different ways. Superficially speaking, some of the participants made a summary of the three articles, whereas others established their own agreement by using materials from the three articles to support their arguments. Second, they made use of the material resources in the articles at different stages of writing. For example, some of them consulted the articles at the drafting stage, and some of them referred to the articles for specific information while writing. They differed in how they orchestrated the activities involved in writing the multiple-source essay. McGinley (1992) examined composing from multiple sources through a series of case studies of readers involved in the process of creating text on the basis of multiple texts. The results showed that the composing process was not a linear one. While students spent much of their time reading the articles at the beginning of the task, they also continued to return to the text and reread throughout the composing process.

The commonality among the empirical studies noted above is that reading the source texts functioned primarily to help them acquire new information about the topic, and fluent and less fluent readers differed in their ability to utilize the sources. The position of these studies is that the writing tasks in these studies were based on the reading, and the nature is linking reading and writing activities thematically may lead to better performance. In extensive reading studies, the writing tasks can hardly be based on any particular text because of the huge reading amount. On the reading and writing relationship, extensive reading studies look at this relationship from a much more global perspective, that is, whether the immersion in large amount of input will benefit writing proficiency.

Extensive reading has proved to be effective on many components of writing. Janopoulos (2009) examines whether sustained silent reading is an effective means of facilitating L2 writing proficiency. He discusses the reading and writing relationship in L2 literacy development, and argues that sustained silent reading can enhance and reinforce many aspects of the learning process for L2 writing proficiency. A three-month extensive reading program was implemented to explore the relationship between extensive reading and readers' linguistic skills. The participants in this study demonstrated remarkable gains in terms of writing proficiency (Tudor & Hafiz, 1989). In terms of the quality of the written output, Lee and Hsu's (2009) study shows the written output produced by readers are superior to the written output produced by non-readers in terms of content, discourse structure, lexical density and language use. Saleem's (2010) study complements Lee and Hsu's by showing that the students' writing proficiency improves in terms of length, organization, and control of genres.

In comparison, findings from the second research mainstream were more conservative on the relationship between extensive reading and writing proficiency. Lightbown et al. (2002) reported on the English skills of students who had learned English as a second language in a three-year experimental comprehension-base program, i.e., a combination of extensive reading activities and extensive listening activities. Participants in the experimental group were offered picture books intended for children who use English as their mother tongue and simplified books written for ESL learners. Each of the books was

accompanied by its corresponding tape recording. The participants read on their own, and used individual tape players and earphones if they wished. Reading and listening were the only activity that the experimental population engaged in their daily English classes. After class, the experimental participants were requested no oral or written production on what they read or listened. After three years of this program, the students in the experimental group performed as well as the students in the comparison group who were taught using the Audio-Lingual Method (see Section 1.2.2 for more information about this method) on measures of listening and reading comprehension and vocabulary recognition. However, students in the comparison group turned out to be better writers than the participants in the experimental group. The explanation of the researchers for this was that the students in the comparison group were able to access their L2 vocabulary more readily when writing because they had had practice in writing short essays in their English classes. In addition to smaller vocabulary size for writing, the participants in the experimental group were found to make more errors in spelling than the students in the comparison group. Although it could be explained as the participants in the experimental group read so well that they did not allocate attention to orthography, spelling error was indeed a defect in written production. Flahive and Bailey (1993) examined a number of first language and second language reading-writing variables in order to examine whether it was the case that better readers tended to be better writers. Their study involved forty university-level non-native speakers of English from twelve different language backgrounds. No relationship was found between the amount of time the participants reported reading for pleasure and scores on a writing tests.

These empirical studies indicate that the amount of pleasure reading is a contributing factor of second language writing ability. Findings from the present study are consistent with this argument. As shown in Table 4.55 below, in the pre-test, the Control group was ahead of the Free Reading group by 0.67 point (CG=11.97, FR=11.30) and the Complementary Reading group by 0.91 (CG=11.97, CR=11.06) point when the full mark of this section was twenty points, but there is no significant main effect for the treatments as a factor in this section as presented in Table 4.55 below. By the mid-test, extensive reading seemed to generate a positive effect on writing as the two experimental groups

demonstrated a considerable advantage to the Control group. In terms of mean score, the Complementary Reading group was ahead of the Control group by 0.08 point (CR=5.67, CG=5.59) and the Free Reading group was ahead of the Control group by 0.52 point (FR=6.11, CG=5.59) when the full mark of the writing section in the mid-test was ten points. What is particularly worth noting is that the difference between the Free Reading group and the Control group was statistically significant ($p=0.04$), as shown in Table 4.56 below.

Table 4. 55 Descriptive statistics and results of ANOVA from the writing sections in the pre-, mid- and post-tests

| Sections (Total score) | Group | N | Mean | SD | F | Sig. |
|------------------------|-------|----|-------|------|------|------|
| Pre-test (20) | CR | 33 | 11.06 | 2.08 | 1.82 | 0.17 |
| | FR | 33 | 11.30 | 2.07 | | |
| | CG | 33 | 11.97 | 1.86 | | |
| Mid-test (10) | CR | 33 | 5.67 | 0.70 | 2.43 | 0.09 |
| | FR | 33 | 6.11 | 1.00 | | |
| | CG | 33 | 5.59 | 1.29 | | |
| Post-test (10) | CR | 33 | 5.57 | 0.73 | 8.19 | 0.00 |
| | FR | 33 | 6.39 | 0.77 | | |
| | CG | 33 | 6.07 | 0.98 | | |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

The situation changed dramatically in the post-test. As shown in Table 4.56 below, the difference between the Complementary Reading group and the Free Reading group was 0.82 point (CR=5.57, FR=6.39), and this difference was statistically significant ($p<0.001$). The difference between the Complementary Reading group and the Control group was 0.50 point (CR=5.57, CG=6.07), and this difference was also statistically significant ($p=0.02$). The gap between the Free Reading group and the Control group was narrowed to 0.32 point compared with the 0.67 point difference between these two groups in the mid-test, and the difference was not statistically significant ($p=0.12$), which means the statistically significant difference between the Free Reading group and the Control group in the mid-test disappeared in the post-test.

The results in the pre, mid- and post-tests revealed that extensive reading gave a good start in improving the participants' writing proficiency, as evidenced by the significant difference between the Free Reading group and the control group in the writing section in the mid-test. However, the participants in the Free Reading group and the control group, who had been explicitly taught to produce accurate

structures and strength the in-class instruction by doing more grammar and translation exercises, outperformed the Complementary Reading group, who devoted less time to grammar instruction in class and did little grammar and translation exercise out of class. This suggests that continued development in English language writing, especially in terms of accuracy, is enhanced by opportunities for output practice as well as explicit instruction from a teacher.

Table 4. 56 The results of multiple comparisons by the Fisher's LSD test in the writing sections in the pre-, mid- and post-tests

| Test (Total score) | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|-----------|-----------|-----------------------|------------|------|
| Pre-test (20) | CR | FR | -0.24 | 0.51 | 0.63 |
| | | CG | -0.33 | 0.51 | 0.51 |
| | FR | CR | 0.24 | 0.51 | 0.63 |
| | | CG | -0.09 | 0.51 | 0.86 |
| | CG | CR | 0.33 | 0.51 | 0.51 |
| | | FR | 0.09 | 0.51 | 0.86 |
| Mid-test (10) | CR | FR | -0.44 | 0.25 | 0.08 |
| | | CG | 0.08 | 0.25 | 0.76 |
| | FR | CR | 0.44 | 0.25 | 0.08 |
| | | CG | 0.52 | 0.25 | 0.04 |
| | CG | CR | -0.08 | 0.25 | 0.76 |
| | | FR | -0.52 | 0.25 | 0.04 |
| Post-test (10) | CR | FR | -0.82 | 0.21 | 0.00 |
| | | CG | -0.50 | 0.21 | 0.02 |
| | FR | CR | 0.82 | 0.21 | 0.00 |
| | | CG | 0.32 | 0.21 | 0.12 |
| | CG | CR | 0.50 | 0.21 | 0.02 |
| | | FR | -0.32 | 0.21 | 0.12 |

* The mean difference is significant at the 0.05 level.

Note: The results of Levene tests show there is no deviation from homogeneity of variance.

To summarize, the position of the present study on the relationship between reading and writing, as supported by the test results in both the translation and writing sections, is that acquisition does occur when learners read for information and general understanding, but it do not lead to significant improvement in learners' ability to produce more accurate structures. This limitation of extensive reading is more acute in the Chinese context, where the dominant English language teaching method, i.e., the Grammar-Translation Method, gives top priority to accuracy and the assessment criteria also place great emphasis on accuracy. There is less evidence that reading assists in writing development in the sense of improving grammatical features.

Another group of researchers show great interest in whether additional writing

activities will enhance the effectiveness of reading programs. The research interest on the relationship between adding writing activities and enhancement of extensive reading comes from the assumption that output reinforces the effect of input (Swain, 1995; Schmidt, 1995). Tsang (1996) compares the effects of reading and writing on writing performance. In this three-month program, the participants (N=144) were equally assigned into three groups of 48. The first group was referred to as a reading group, the second group a writing group and the third group a math group. What the three groups shared in common was that all the three groups adhered to normal English curriculum, and the difference between them, as the groups names suggested, was the assignment. The reading group was required to read eight graded readers and wrote up one book review for each graded reader. The writing group was required to write eight essays on eight given topics covering a variety of content and genres. The math group served as a control group, and the purpose of the math assignment was to ensure the extracurricular exercise did not provide English input. At the end of the program, the three groups were given thirty minutes to write an essay on the same topic. Results showed there was a significant main effect for the assignment as a factor in the participants' writing. The reading group significantly outperformed the writing group and the math group, and there was significant difference between the writing group and math group on writing.

Mason (2004) examines whether adding supplementary post-reading activities in an extensive reading program would enhance the effect of extensive reading on developing grammar accuracy. Participants (N=104) in this study were EFL college learners in Japan, and they were stratified randomly into three groups. Each participant read 500,000 words on average in three academic terms. Participants in the first group wrote summaries of what they read in their first language, Japanese, and later this group will be referred to as Japanese summary group. The participants in the other two groups wrote summaries wrote summaries of what they read in English. The key difference between the latter two groups was that one group did not receive corrective feedbacks (referred to as English summary group), and one group not only received corrective feedbacks but also rewrote their summaries in light of the feedbacks (referred to as correction group). Three measures were used to collect data: a 100-item cloze

test, the reading section of the TOEIC test and the number of error free clauses made per 100 words in writing. Data analysis showed there was no significant main effect for the treatments as a factor on the test results. It could be concluded that adding supplementary writing did not result in greater accuracy.

Findings from the present study are consistent with those of Tsang's and Mason's studies. The complementary reading participants wrote a brief report every day. The researcher read the reports in order to ensure they had actually read the books, but did not mark them or correct the spelling, content and grammar mistakes. Neither did she blame the students whose reports were of low writing quality. The test results of the writing section in the post-test indicate that the Complementary Reading Group was the weakest in writing among the three groups, and there were significant differences between the complementary reading group and the other two groups.

The limitation of the present study in exploring the relationship between extensive reading and writing proficiency is that it is unknown in which aspect of writing the complementary reading participants were weak. When their compositions were marked, the markers keyed in a single mark for students' writing without further indicating the marks for separate components, such as lexical accuracy and density, grammatical accuracy, cohesion and coherence, etc, were not done, and thus the participants' performance on these subscales were not able to be analyzed and compared. What we can infer from the test results is that the Complementary Group did not do better in this section because of their daily writing activities. And this study complements Mason's and Tsang's studies that output practice, with or without correction, does not enhance the effect of comprehensible input. At this stage it can be concluded that additional writing activities do not enhance extensive reading.

4.7 Findings from the delayed post-test

The primary focus of the delayed post-test was to collect data to examine the longitudinal effect of extensive reading. As shown in Table 4.57 below, the situation changed in the delayed post-test. By the time of the post-test, statistically significant differences were detected in listening, cloze, reading,

translation, writing and total score (see Table 4.14 in Section 4.2 for detailed statistics of the post-test). In comparison, statistically significant differences continued to be found in translation, writing and total score in the delayed post-test. This is indicative of the students' further development in the productive skills one year after the program ended. The results also suggest that it takes a longer time for the effect of extensive reading to emerge in the domain of productive skills compared to receptive skills.

Table 4. 57 Descriptive statistics and results of ANOVA from the delayed-test

| Section (Total score) | Group | N | Mean | SD | Std. Error | F | Sig. |
|-----------------------|-----------------------|----|--------|-------|------------|------|------|
| Listening (30) | CR | 32 | 22.38 | 2.21 | 0.39 | 1.93 | 0.15 |
| | FR | 33 | 22.67 | 2.78 | 0.48 | | |
| | CG | 32 | 21.41 | 3.04 | 0.54 | | |
| Grammar (16) | CR | 32 | 12.41 | 2.00 | 0.35 | 1.89 | 0.16 |
| | FR | 33 | 12.30 | 1.94 | 0.34 | | |
| | CG | 32 | 11.53 | 1.97 | 0.35 | | |
| Vocabulary (9) | CR | 32 | 6.19 | 1.51 | 0.27 | 0.00 | 1.00 |
| | FR | 33 | 6.21 | 1.11 | 0.19 | | |
| | CG | 32 | 6.22 | 1.39 | 0.24 | | |
| Cloze (15) | CR | 32 | 11.31 | 1.79 | 0.32 | 2.29 | 0.11 |
| | FR | 33 | 10.97 | 1.67 | 0.29 | | |
| | CG | 32 | 10.38 | 1.86 | 0.33 | | |
| Reading 1(27) | CR | 32 | 21.28 | 2.68 | 0.47 | 0.20 | 0.82 |
| | FR | 33 | 21.09 | 3.47 | 0.60 | | |
| | CG | 32 | 20.75 | 3.88 | 0.69 | | |
| Reading 2 (8) | CR | 32 | 5.20 | 1.65 | 0.29 | 2.06 | 0.13 |
| | FR | 33 | 5.88 | 1.63 | 0.28 | | |
| | CG | 32 | 5.19 | 1.42 | 0.25 | | |
| Translation (20) | CR | 32 | 14.11 | 2.05 | 0.36 | 4.65 | 0.01 |
| | FR | 33 | 14.41 | 1.83 | 0.32 | | |
| | CG | 32 | 13.04 | 1.78 | 0.31 | | |
| Writing (25) | Content (10) | CR | 32 | 5.86 | 0.64 | 5.27 | 0.01 |
| | | FR | 33 | 6.23 | 0.63 | | |
| | | CG | 32 | 5.70 | 0.74 | | |
| | Language quality (10) | CR | 32 | 5.70 | 0.66 | 4.95 | 0.01 |
| | | FR | 33 | 6.08 | 0.70 | | |
| | | CG | 32 | 5.53 | 0.78 | | |
| | Organization (5) | CR | 32 | 1.81 | 0.45 | 0.22 | 0.80 |
| | | FR | 33 | 1.89 | 0.45 | | |
| | | CG | 32 | 1.84 | 0.57 | | |
| | Total (25) | CR | 32 | 13.38 | 1.40 | 4.35 | 0.02 |
| | | FR | 33 | 14.20 | 1.48 | | |
| | | CG | 32 | 13.08 | 1.85 | | |
| Total (150) | CR | 32 | 106.25 | 8.05 | 1.42 | 4.41 | 0.01 |
| | FR | 33 | 107.72 | 8.43 | 1.47 | | |
| | CG | 32 | 101.59 | 9.45 | 1.67 | | |

* The mean difference is significant at the 0.05 level.

Note 1: The results of Levene tests show there is no deviation from homogeneity of variance.

Note 2: Reading 1 was in the form of multiple choice and Reading 2 was short questions and answers.

A review of the test results of the translation and writing sections in the pre-,

mid- and post-tests (see Table 4.14 in Section 4.2 for the detailed statistics of the post-test) shows that the Free Reading treatment gave the participants a good start and an early advantage in improving their writing proficiency, as evidenced by the significant difference between the Free Reading group and the control group in the writing section in the mid-test (FR=6.11, CG=5.59, $p=0.04$). This reveals that extensive reading plus explicit instruction are important. The results of the present study show that the benefits of extensive reading are most visible in the Free Reading treatment, which integrated extensive reading and explicit instruction.

The advantage produced by the integration of extensive reading and explicit instruction does not deny the effect of extensive reading alone on improving students' writing proficiency. As shown in Table 4.58 below, the Complementary Reading group outperformed the Control group in terms of content, language quality and total score in the writing section in the delayed post-test, although the differences in these comparisons were not statistically significant. It would be interesting to see if a delayed post-test at a later time, say in another six months or one year, would yield a statistically significant difference; longitudinal studies with longer time spans in future research would be valuable to determine the effects of extensive reading treatments over time.

It is worth noting that the Complementary Reading group also significantly outperformed the Control group in the translation section in the delayed post-test (CR=14.11, CG=13.04, $p=0.03$).

It is possible that the Complementary Reading treatment, although it was not supplemented by explicit instruction, may have contributed to the students' desired (albeit delayed) improvement in writing proficiency. However, attention should be drawn to another possibility as well, namely, that the Complementary Reading group reverted to the normal curriculum after the extensive reading treatment ended, and hence the participants may have benefited from the combined effects of extensive reading plus subsequent explicit instruction.

Table 4. 58 The results of multiple comparisons by the Fisher's LSD test in the delayed post-test

| Dependent Variable | (I) Group | (J) Group | Mean Difference (I-J) | Std. Error | Sig. |
|--------------------|------------------|--------------|--------------------------|------------|------|
| Listening | CR | FR | -0.29 | 0.67 | 0.66 |
| | | CG | 0.97 | 0.67 | 0.15 |
| | FR | CR | 0.29 | 0.67 | 0.66 |
| | | CG | 1.26 | 0.67 | 0.06 |
| | CG | CR | -0.97 | 0.67 | 0.15 |
| | | FR | -1.26 | 0.67 | 0.06 |
| Grammar | CR | FR | 0.10 | 0.49 | 0.83 |
| | | CG | 0.88 | 0.49 | 0.08 |
| | FR | CR | -0.10 | 0.49 | 0.83 |
| | | CG | 0.77 | 0.49 | 0.12 |
| | CG | CR | -0.88 | 0.49 | 0.08 |
| | | FR | -0.77 | 0.49 | 0.12 |
| Vocabulary | CR | FR | -0.02 | 0.33 | 0.94 |
| | | CG | -0.03 | 0.34 | 0.93 |
| | FR | CR | 0.02 | 0.33 | 0.94 |
| | | CG | -0.01 | 0.33 | 0.98 |
| | CG | CR | 0.03 | 0.34 | 0.93 |
| | | FR | 0.01 | 0.33 | 0.98 |
| Cloze | CR | FR | 0.34 | 0.44 | 0.44 |
| | | CG | 0.94 | 0.44 | 0.04 |
| | FR | CR | -0.34 | 0.44 | 0.44 |
| | | CG | 0.59 | 0.44 | 0.18 |
| | CG | CR | -0.94 | 0.44 | 0.04 |
| | | FR | -0.59 | 0.44 | 0.18 |
| Reading 1 | CR | FR | 0.19 | 0.84 | 0.82 |
| | | CG | 0.53 | 0.84 | 0.53 |
| | FR | CR | -0.19 | 0.84 | 0.82 |
| | | CG | 0.34 | 0.84 | 0.69 |
| | CG | CR | -0.53 | 0.84 | 0.53 |
| | | FR | -0.34 | 0.84 | 0.69 |
| Reading 2 | CR | FR | -0.68 | 0.39 | 0.09 |
| | | CG | 0.02 | 0.39 | 0.97 |
| | FR | CR | 0.68 | 0.39 | 0.09 |
| | | CG | 0.69 | 0.39 | 0.08 |
| | CG | CR | -0.02 | 0.39 | 0.97 |
| | | FR | -0.69 | 0.39 | 0.08 |
| Translation | CR | FR | -0.30 | 0.47 | 0.53 |
| | | CG | 1.07 | 0.47 | 0.03 |
| | FR | CR | 0.30 | 0.47 | 0.53 |
| | | CG | 1.36 | 0.47 | 0.00 |
| | CG | CR | -1.07 | 0.47 | 0.03 |
| | | FR | -1.36 | 0.47 | 0.00 |
| Writing | Content | FR | -0.37 | 0.17 | 0.03 |
| | | CG | 0.16 | 0.17 | 0.35 |
| | | CR | 0.37 | 0.17 | 0.03 |
| | | CG | 0.52 | 0.17 | 0.00 |
| | | CR | -0.16 | 0.17 | 0.35 |
| | | FR | -0.52 | 0.17 | 0.00 |
| | Language Quality | FR | -0.37 | 0.18 | 0.04 |
| | | CG | 0.17 | 0.18 | 0.34 |
| | | CR | 0.37 | 0.18 | 0.04 |
| | | CG | 0.54 | 0.18 | 0.00 |
| | | CR | -0.17 | 0.18 | 0.34 |
| | | FR | -0.54 | 0.18 | 0.00 |
| | Organization | FR | -0.08 | 0.12 | 0.51 |
| | | CG | -0.03 | 0.12 | 0.80 |
| | | CR | 0.08 | 0.12 | 0.51 |
| | | CG | 0.05 | 0.12 | 0.68 |
| | | CR | 0.03 | 0.12 | 0.80 |
| | | FR | -0.05 | 0.12 | 0.68 |
| | Total | FR | -0.82 | 0.39 | 0.04 |
| | | CG | 0.30 | 0.40 | 0.46 |
| | | CR | 0.82 | 0.39 | 0.04 |
| | | CG | 1.12 | 0.39 | 0.01 |
| | | CR | -0.30 | 0.40 | 0.46 |
| | | FR | -1.12 | 0.39 | 0.01 |
| Total | CR | FR | -1.47 | 2.15 | 0.49 |
| | | CG | 4.66 | 2.17 | 0.03 |
| | FR | CR | 1.47 | 2.15 | 0.49 |
| | | CG | 6.13 | 2.15 | 0.01 |
| | CG | CR | -4.66 | 2.17 | 0.03 |
| | | FR | -6.13 | 2.15 | 0.01 |

* The mean difference is significant at the 0.05 level.

Note 1: The results of Levene tests show there is no deviation from homogeneity of variance.
Note 2: Reading 1 was in the form of multiple choice and Reading 2 was short questions and answers.

Last but more importantly, although the two extensive reading treatments have their own strong points in improving students' writing proficiency, we advocate for the necessity to integrate extensive reading and explicit instruction in building the students' writing proficiency. As shown by the test results in the pre-, mid-, and post-tests, the Free Reading performed the best in the translation and writing sections in these three tests. Further, in the writing section in the delayed post-test the Free Reading group continued to significantly outperform the Control group (FR=14.20, CG=13.08, $p=0.01$), and more impressively, still outperformed the Complementary Reading group (FR=14.20, CR=13.38, $p=0.04$). This argues strongly for the power of the Free Reading treatment, i.e. simultaneous extensive reading plus explicit instruction.

4.8 Findings from the reading records

4.8.1 Reading speed

For many ESL and EFL readers, reading in English is a slow and arduous process (Hamp-Lyons 1983; Cooper 1984); yet fast reading does not receive much attention from either teachers or students in the Chinese context. The first reason for neglecting rapid reading is that the reading comprehension section in the NMET is composed of three to four short passages with 300 to 500 words (see Section 1.4 for detailed introduction of the NMET). Because of the signal from the highest-stakes English exam, it was likely that students equated reading in English with reading short passages, which naturally kept them away from developing reading speed. The second reason is that English language teaching in classroom is mostly based on short texts (see Section 3.6 for the detailed introduction of the textbook used in the present study, *Oxford English for Senior High School*). In traditional English lessons, texts are used for the presentation of language points, either grammar rules or lexical items, rather than for reading purposes.

Although the above two reasons may explain the absence of fast reading training in traditional English curriculum in mainland China, the value of fast reading is

supported by research literature and development in reading speed is strongly recommended by researchers (e.g., Fry, 1963; Carver, 1992). The vicious circle of the weak reader presented in Nuttall (1996) is that the weak readers do not understand what they read because of lack of vocabulary, so they read slowly. They do not enjoy reading. Consequently they do not read much. As a result the vicious circle repeats. One of the direct consequences of slow reading is that the readers will be exposed to much less amount of comprehensible input, which is central to language acquisition.

Despite the shared concern on the value of fast reading, there is disagreement on the optimal or standard reading rate. The optimal reading speed ranges from 180 words per minute (wpm) (Higgins & Wallace, 1989), to 300 wpm (Jenson, 1986). Unlike the unified standard for all readers noted above, Fry (1975) suggests a flexible setup for readers of different levels of reading proficiency: 350 wpm for competent readers, 250 wpm for average readers and 150 wpm for poor readers.

The various optimal reading speeds suggested by the research literature ranges from 180 wpm minute to 300 wpm. These suggested optimal reading speeds result in two problems. The first problem is that the above-noted optimal reading speeds are suggested for readers of a certain age. We cannot deny that the older the readers are, the more likely they are to read faster. But how fast should ESL and EFL secondary students read at their own ages on average? For this, Anderson consulted the national silent reading rates by grade level (Hasbrouck & Tindal, 2006, cited in Heidi, 2010).

Silent Reading Rates

1st grade: 80 wpm

2nd grade: 115 wpm

3rd grade: 138 wpm

4th grade: 158 wpm

5th grade: 173 wpm

6th grade: 185 wpm

7th grade: 195 wpm

8th grade: 204 wpm

9th grade: 214 wpm

10th grade: 224 wpm

11th grade: 237 wpm

12th grade: 250 wpm

College or University: 280 wpm

This framework reveals the gap between the Chinese EFL learners' reading speed and the proposed standard. As specified in the Chinese English Curriculum Standards, the senior high school students should be able to read at 60 to 80 wpm, and the requirement for college and university students is 100 wpm. According to the above standard, the Chinese senior high school and university students are not even able to read half as fast as their international peers. It is not a problem caused by the underestimation of students' reading speed, but in fact the Board of Education does not make ambitious demands on students' reading in English.

As illustrated in Table 4.59 below, the English Curriculum Standards (Shanghai version) specifies the amount of extracurricular reading. The total reading amount for 12 years is 600,000 words at most, which is equal to 137 words per day. In particular, at the senior high school level, a senior high school student in Shanghai is required to read 250,000 running words in three years. A simple calculation tells us that a student needs to read about 230 words per day if he or she keeps reading every day. If he reads every other day, 500 words each time is enough. Under such a circumstance the students will not feel any need to read fast. The washback effect of the National Matriculation English Test (NMET) also contributes to the students' low reading speed. From the students' perception of this highest-stakes English test in mainland China, the texts used for assessing

reading ability are short (i.e., 300 to 500 words). It is likely for them to equate reading in English to reading short passages, which naturally keep them away from fast reading.

Table 4. 59 Amount of extracurricular reading required in the National English Curriculum Standard

| School stage | Level | Amount of extracurricular reading (word) | |
|--------------------|----------|--|---------------------|
| | | Net | Accumulative |
| Primary school | Basic | 800,000 | 800,000 |
| | Advanced | 200,000 | 1,000,000 |
| Junior high school | Basic | 1,800,000 | 2,600,000-2,800,000 |
| | Advanced | 200,000 | 2,800,000-3,000,000 |
| Senior high school | Basic | 2,500,000 | 5,100,000-5,500,000 |
| | Advanced | 500,000 | 5,600,000-6,000,000 |

The second problem is that students are hardly able to read materials of different levels of difficulty at the same speed. How fast should they read when the levels of difficulty of the reading materials vary? There is no doubt that they will vary their reading speeds when reading at different levels of difficulty, but we do not know how big the change will be.

The gap between the research literature and reality suggests a number of areas of enquiry for this study. Our first focus is to investigate how fast the Chinese senior high school students can read, especially when the levels of the difficulty of the reading materials vary. The second focus is to gain insight into the effect of extensive reading on promoting students' reading speed.

The following two areas of enquires are to be addressed in the subsequent discussion:

1. How fast can the Senior One students in Chinese mainland read? How fast can they read when the difficulty levels of the reading materials vary?
2. What mode of change in reading speed will the two groups of extensive readers show with extensive reading going on?

The change in students' reading speed was recorded in their individual reading records. The information provided by the reading record includes: the participants' reading speeds for the first Level 1 book and the last Level 1 book, the reading speeds for the first Level 2 and the last Level 2 book, the reading

speeds for the first Level 3 and the last Level 3 book, and the reading speeds for the first Level 4 and the last Level 4 book. The data presented in Table 4.60 below are the averages of the reading speeds of the 33 participants in each group. Not all the students had read the Level 5 and Level 6 books so this part of the data are not presented here, as the purpose of our discussion was to detect something in common shared by all the participants.

Table 4. 60 Reading speeds at different levels of the two experimental groups

| Group | Le1-F | Le1-L | Le2-F | Le2-L | Le3-F | Le3-L | Le4-F | Le4-L |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|
| CR | 139.89 | 183.12 | 134.59 | 191.48 | 159.97 | 190.13 | 178.58 | 230.66 |
| FR | 145.13 | 145.83 | 139.52 | 175.39 | 167.46 | 173.87 | 205.42 | 212.18 |

Note: Le1-F=the first Level 1 book, Le1-L=the last Level1 book, etc.

According to the standards proposed by Hasbrouck and Tindal (2006) which is noted above, the Chinese EFL students read at a rather low speed. The suggested optimal reading speed for the 10th grader, namely senior one, is 224.00 wpm. In contrast, when the present study began, the average reading speed of the complementary and Free Reading group was 139.89 wpm and 145.14 wpm respectively, less than 65% of the international standard. Given that the average speed was based on the fact that the participants read at Level 1, namely the 400 headword level, at which they read without any difficulty, the situation seemed more unpromising. When the Complementary Reading group moved to L2, their average reading speed reached 183.12 wpm, an approximately 30% increase. In comparison, the average reading speed of the Free Reading participants remained stable. After reading seven Level books on average, there was no obvious progress in their reading speed.

When moving to L2, both the Complementary Reading group and Free Reading group showed retrogression in reading speed. The Complementary Reading group stepped backward from 183.12 wpm to 134.59 wpm, and the Free Reading group retreated from 145.83 wpm to 139.52 wpm. The two groups read slower than their beginning reading speed when the program began. It can be seen that the difficulty level of reading materials did have an impact on reading speed. However, the impact did not last long. When the participants moved to Level 3, both groups reached a new peak: 191.48 wpm for the Complementary Reading group and 175.39 wpm for the Free Reading group. In the subsequent phase of

the program, the same situation repeated each time the participants moved from a lower level to a higher level. Accompanying each move there was always retrogression in reading speed, and there was always a new peak when the participants finished one level. We have the reason to believe that the varying levels of difficulty of the reading materials did impact reading speed, but the impact was temporary as the participants regained their reading speed once they were used to the new difficulty level. In addition, they made progress in their reading speed as they continued reading.

At the end of the program, the complementary reading participants were able to read at 230.66 wpm, faster than the suggested 224.00 wpm. The free reading participants, although slower than their complementary reading peers, could manage 212.18 wpm, slightly slower than the proposed standard. In the whole program, the complementary reading participants showed a 64.8% increase in their reading speed, from 139.89 wpm to 230.66 wpm. The free reading participants moved from 145.13 wpm to 212.18, a 46.20% increase. We should also be reminded that the increase was based on their move from Level 1 to Level 4, from the 400 headword level to the 1,400 headword level. It is reasonable to assume that the increase in their reading speed is greater than the simple calculation shows. Conservatively estimated, the increase in the Complementary Reading group's reading speed should be around 70% to 100%, and the estimation for the Free Reading group should be around 50% to 80%.

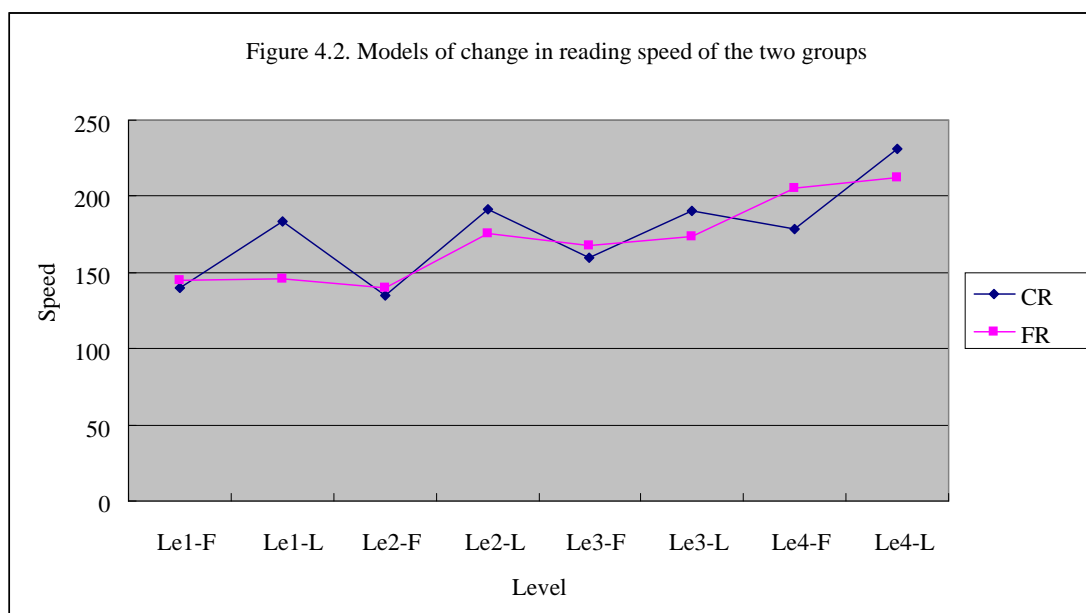
A further calculation showed where and when the biggest progress happened. As shown in Table 4.61, the Complementary Reading group experienced rapid growth in their reading speed. The increase within the same level ranged from 18.85% to 42.27%, and the total increase rate in the whole program was 64.89%. In comparison, the biggest progress of the Free Reading group occurred during the L2 reading period. .

Figure 4.2 illustrated that the modes of the two groups' progress in reading speed were different. Throughout the whole program, the Free Reading group generally made steady and gradual progress as they moved upward. There was no obvious retrogression when the free reading participants moved from a lower level to a

Table 4. 61 Increase in reading speed between each two levels of the two experimental groups

| Group | Le1-F | Le1-L | % | Le2-F | Le2-L | % | Le3-F | Le3-L | % | Le4-F | Le4-L | % | Total |
|-------|--------|--------|-------|--------|--------|-------|--------|--------|-------|--------|--------|-------|-------|
| CR | 139.89 | 183.12 | 30.90 | 134.59 | 191.48 | 42.27 | 159.97 | 190.13 | 18.85 | 178.58 | 230.66 | 29.16 | 64.89 |
| FR | 145.13 | 145.83 | 0.004 | 139.52 | 175.39 | 25.71 | 167.46 | 173.87 | 3.83 | 205.42 | 212.18 | 3.29 | 46.20 |

higher level. Compared with the relatively smooth progress of the Free Reading group, the curve of the Complementary Reading group fluctuates. Within each level, the Complementary Reading group made notable progress; however, there must be retrogression in reading speed when they began a higher level.



The different trends of the two groups shown by the curves reveal the following implications. First of all, the Complementary Reading participants were not able to choose what they felt like reading; instead, the individual Complementary Reading participant had to keep up with the pace of the whole group. They improved their reading speed at a certain level because of the huge reading amount; however, when requested to move to a higher level, not all the participants were well prepared for the move. That was why without exception there was obvious retrogression when the Complementary Reading group began a new level. In contrast, the Free Reading participants were able to manage their own reading. They decided what and how much they read. The Individual Reading records showed that the Free Reading participants did not follow the levels strictly; instead, their choice of books was largely because of their personal taste. A reasonable explanation was that the Free Reading participants had a clear perception of the difficulty levels of the OBW graded readers so they adjusted their reading speed, skills and strategies to cater for the change. It must be pointed out that 250 headword, 400 headword, 700 headword, 1000 headword

and 1400 headword required different decoding efforts although the pre-test showed that the 1400 headword was well within the participants' reading competence. Because of their proper management of their reading, the Free Reading participants made steady increase by 46.20%, although their progress between each two levels was less noticeable than that of their Complementary Reading peers.

In addition to showing how fast the two groups of participants read, the statistics describing reading speed reveals more than it is supposed to. Grabe (2009) lists six academic purposes for reading:

1. Reading to search for information (scanning and skimming)
2. Reading for quick understanding (skimming)
3. Reading to learn
4. Reading to integrate information
5. Reading to evaluate, critique, and use information
6. Reading for general comprehension (in many cases, reading for interest or reading to entertain)

(Grabe, 2009, p.8)

Among the six purposes, the first two are carried out at very high speed. In comparison, reading to learn places more processing demands on the readers because the reader is expected to remember the main ideas and many supporting ideas and be able to recall this information as needed (Enright et al., 2000). Based on this, it is understandable that we usually read to learn at a relatively lower speed, i.e. 200 wpm in Carver (1992). The fact that reading to learn is carried out at a relatively lower speed does not deny the possibility that the participants can also learn when they read faster for general comprehension, however, one cannot guarantee that learning necessarily take place.

If our interpretation of the statistics describing how the two groups read differently is valid, namely the Complementary Reading group read for quick understanding (skimming) and the Free Reading group read to learn and also for pleasure, it partly explains why the Complementary Reading group read five times as much as the Free Reading group but did not perform as well as the Free

Reading group in the post-test. There are a number of studies in educational psychology which demonstrate significant differences in comprehension processing as a result of different purposes for reading (e.g., Linderholm & van den Broek, 2002; Carver, 1992). These studies demonstrate that the readers adopt different processing goals when they read for different purposes. In the present study, the Complementary Reading participants were required to read about 5,000 words per day, probably sometimes at a *i*+1 level or a level they were not comfortable with. What was worse, they were faced with heavy academic pressure, i.e., a lot of daily assignment, as described in Chapter 3. In this situation, it is entirely understandable, as well as inevitable, that the Complementary Reading participants read for quick understanding, which would enable them to finish their daily book report. In comparison, the Free Reading participants read in a much more relaxing environment from the start to the end of this program. They participated in this program on a voluntary basis, which indicated their more positive attitude towards reading in English when the present study began. They were free to choose what they want to read, and they were able to read at levels that they were comfortable with. All these favorable conditions for the Free Reading participants enabled them to read for study purpose, in which significantly more strategies and knowledge were involved, and eventually contributed to their remarkable achievement in the post-test.

Another important point that must be reiterated here is that reading to learn, which is closely associated with a relatively lower reading speed, does not mean reading fast or reading for quick understanding is without any merit. The Complementary Reading groups' achievement in EPER levels, as measured by the EPER test, demonstrated that the Complementary Reading participants become fluent readers in English. They gained confidence in reading in English, and are expected to be able to read authentic unsimplified reading materials after they enroll in higher education institutions.

Conclusion

A gradual increase in students' reading speed is noted, at both the individual level and the whole group level. The Complementary reading participants began the program with a speed at 139.89 wpm for Level 1 books on average and ended at

a speed at 230.66 wpm for Level 4 books. The Free Reading participants began the program with a speed of 145.13 wpm for Level 1 books on average and ended at a speed at 212.18 wpm for Level 4 books. On individual levels, participants in both groups read the last book faster than the first book at the same level. The Complementary Reading participants made bigger progress in this aspect, since their increase in reading speed on individual levels was bigger than that of the Free Reading group. It could be attributed to their huge reading quantity and daily practice.

In addition to the progress at individual levels, students in both groups made remarkable progress as extensive readers. Judging from the absolute number, the average reading speed of the students in the Complementary Reading group increased by about 64.89%, from 139.89 wpm to 230.66 wpm. The increase for the Free Reading group was 46.20%, from 145.13 wpm to 212.18 wpm. Furthermore we need to note that OBW Level 1 is at 400-headword level while OBW Level 4 is at 1400-headword level. Taking the increasing levels of difficulty of reading materials into account, students' progress in reading speed was much greater than what the absolute numbers had shown.

Despite the progress, the two groups showed different modes of progress mainly because of the treatments. The pedagogical implication drawn from the difference was that extensive reading was effective in improving readers' reading speed but it is advisable not to take the effectiveness for granted. The curves of the two groups showed that a large reading amount would lead to improvement in reading speed, but the readers' self-preparation and self-adjustment would promise a smoother transfer from a lower level to a higher level.

4.8.2 Preference for different genres

The purpose of the discussion in this section is to identify the gap between the academic circle and the teenager readers in terms of their perceptions of what a good graded reader is. The Extensive Reading Foundation (ERF) (<http://www.erfoundation.org>) awards and promotes excellent graded readers for extensive reading every year. In 2010, the book in the OBW series *Playing with Fire: Stories from the Pacific Rim* (Level 3) won the award. The reason was: *The*

stories are all very strong, heart-hitting and well expressed; excellent use of illustrations by local illustrators (<http://www.erfoundation.org/erf/node/67>). In 2011 two books in the finalist were: *The Everest Story* (OBW Level 3), and *Leaving No Footprint: Stories from South Asia* (OBW Level 3).

In addition to the annual book award scheme, recently ERF announces the launch of a new award scheme named in honor of John Milne, the founder of the Heinemann Guided Readers series. Milne believed that text modification in terms of lexis and syntax was not enough to make a book suitable for language learners. He introduced the notion of information control into graded readers. This innovation is crucial in the development of language learner literature, and becomes one of the three components of defining graded readers (see Section 2.3 for detailed discussion of graded readers). The first John Milne Award was presented to David Hill in the first Extensive Reading World Congress in Kyoto, for his remarkable contribution to graded readers and extensive reading research.

The efforts made show that the publishers have been trying to present the readers with high-quality books, and the academic circle has been renewing the standards of good readers. However, the perception of what a good book is of the teenager readers in the present study seem to differ from that of the researchers and publishers.

According to the rating of the 235 titles in the OBW series by the 33 Complementary Reading participants, the rating of the 2010 award-winning book *Playing with Fire: Stories from the Pacific Rim* was 3.44. The nominations for the 2011 award, *The Everest Story* is a true story and *Leaving No Footprint: Stories from South Asia* belongs to World Stories, was rated 3.53 and 4.07 respectively, which were not the best-received by the Complementary Reading participants.

These three cases suggest that the teenager readers' perception of what a good book is differs from that of the professional. A proposed area of inquiry for future research is whether the quality of the reading materials will influence language acquisition through extensive reading. To be specific, this inquiry examines

whether the reading quality will be lowered if the readers miss the best titles in the series, and in turn interfere with the results that extensive reading would produce.

Another observation is that factfiles, world stories and true stories account for a considerable proportion of the least well-received books among the readers. Taking the five least well-received books from Level 1 to 4 as an example (see Table 4.62 below), *World Stories* accounted for 20.0%, *True Stories* comprised 15% and factfiles occupied 15% of the twenty titles respectively. As introduced in Chapter 3, the 235 titles in the OBW series are distributed to nine genres, with 36 *Classics*, 24 *Crime & Mystery*, 24 *Factfile*, 31 *Fantasy and Horror*, 44 *Human Interest*, 10 *Playscripts*, 36 *Thriller and Adventure*, 18 *True Stories* and 12 *World Stories*. In comparison, *World Stories* accounted for 5% in the whole series but 20% in the least well-received titles, *True Stories* comprised 7.7% of the whole series but 15% in the least well-received titles, and *Factfile* occupied 10.2% but 15% in the least well-received titles.

Table 4. 62 The least well-received five titles for Level 1, 2, 3 and 4

| Level | Name | Genre | Rate |
|-------|--|--------------------|------|
| 1 | One-Way Ticket - Short Stories | Human Interest | 3.45 |
| | Ned Kelly: A True Story | True Stories | 3.43 |
| | The Meaning of Gifts - Stories from Turkey | World Stories | 3.42 |
| | The Coldest Place on Earth | True Stories | 3.34 |
| | Mutiny on the Bounty | True Stories | 3.33 |
| 2 | Return to Earth | Fantasy/Horror | 3.35 |
| | Matty Doolin | Human Interest | 3.21 |
| | The Importance of Being Earnest | Playscripts | 3.21 |
| | The Piano | Human Interest | 3.19 |
| | The Pit and the Pendulum and Other Stories | Fantasy/Horror | 3.05 |
| 3 | Wyatt's Hurricane | Thriller/Adventure | 3.36 |
| | Australia and New Zealand Factfile | Factfiles | 3.34 |
| | Who, Sir? Me, Sir? | Human Interest | 3.24 |
| | A Cup of Kindness: Stories from Scotland | World Stories | 3.19 |
| | Tooth and Claw - Short Stories | Human Interest | 2.61 |
| 4 | The History of the English Language Factfile | Factfiles | 3.49 |
| | Great Crimes Factfile | Factfiles | 3.47 |
| | Silas Marner | Classics | 3.47 |
| | Land of My Childhood - Stories from South Asia | World Stories | 3.45 |
| | The Price of Peace: Stories from Africa | World Stories | 3.00 |

The possible explanation for the unpopularity of *World Stories* and *True Stories* is that the stories in these two genres are set in different social and cultural contexts, which may not be familiar to Chinese secondary school students. The

potential threat to *Factfile* is that there is an overload of information in these titles, which would make it difficult for students to process and summarize. The aim of this preliminary and visual comparison is to suggest that background knowledge and information control are two areas for further enquiry in subsequent extensive reading programs.

5 Conclusion and recommendations

5.1 Summary

In the past decades since 1978, the syllabus designers, teachers and students have been trying their best to improve students' English language proficiency, as evidenced by the incorporation of the latest development in linguistics into the national and local syllabuses and the introduction of new methodologies into the classroom. Various initiatives were made to reform English language teaching and learning, but the results have not been encouraging. Therefore it is assumed that there might be some missing links between English language teaching and the learning process.

Since English language teaching in mainland China is constrained by many contextual factors, including the uniformed curriculum standard, the washback effect of the NMET, limited instructional time, rigid curriculum, etc, it is assumed that lack of input is the most important variable that may account for the lower success of English language teaching and learning, and extensive reading is an optimal means of providing comprehensible input. The research literature has supported that extensive reading can effectively improve ESL or EFL learners' English language proficiency, and it is not in conflict with the contextual constraints noted above.

In the present study, the scope was delimited to the students in a senior high school in Shanghai. However, its influence might apply to other parts of China, because the vast majority of senior high schools have undergone the same problems in English teaching and learning.

The major purposes of this research undertaking were to examine the effectiveness of providing input by means of extensive reading in senior high schools in Shanghai and to provide the research findings as reference to recommend some feasible changes in English language teaching and learning in senior high schools.

The present study focused on the following research questions:

1. Does extensive reading result in improvement in Shanghai secondary school students' English language proficiency? Specifically, does extensive reading positively influence the students in terms of language knowledge, namely, grammar and vocabulary, receptive skills, namely reading and listening, or productive skills, namely writing (and translation in the present study)? If so, in what ways?
2. How can extensive reading be more successfully implemented in the Chinese context? Which type of extensive reading program is more effective, one that is integrated as part of the class curriculum, or one that is an extracurricular activity?
3. Which form of extensive reading implementation may be more effective in enhancing a particular skill?

The study adopted one kind of true experimental design called the Pre-test/Post-test Control Groups Design. At the beginning of the program, the participants were stratified by the following criteria: their performance on the pre-test into three groups: the Complementary Reading group, the Free Reading group and the Control group. Each group was composed of 23 girls and 10 boys. All the three groups scored 53.12 points on the pre-test, which showed the three groups had almost equal English language proficiency at the beginning of the present study. Some of the most important extraneous variables were under control during the program, such as the participants' previous English language proficiency, gender, teachers and the amount of time the participants spent on English language learning. The purpose of controlling the extraneous variables was to create almost equivalent teaching and learning conditions for the three groups in order to eliminate or minimize their influence on the results. The independent variable was extensive reading. The dependent variable was the participants' performance on the post-test and the EPER Placement Test B, which all the participants took in June, 2011 after one year of extensive reading. A delayed post-test was administered one year after the program ended, assessing the long-term effect of effect of extensive reading on the participants' English language proficiency.

The following major findings emerged from the data analysis:

1. The results of the locally developed pre- and post-tests and the EPER Placement Test suggest that extensive reading as both a medium of additional input for the FR group and as a component complementary to normal English language curriculum for the CR group result in improvement in the sample students' overall English language proficiency.
2. The post-test results show improvement in specific proficiency areas (i.e. listening, reading, and writing proficiency) and language knowledge components (i.e. grammar and vocabulary). The delayed post-test results show students' further improvement in writing proficiency one year after the program ended; this suggests that it takes more time for extensive reading to show positive effect on the productive skills.
3. The analysis of the reading records suggest that extensive reading significantly improves students' reading speed.
4. Despite its benefits, extensive reading has limitations. It was less effective in helping students acquire some types of vocabulary knowledge, such as low frequency words. It was also less effective in helping students to improve their writing accuracy.

5.2 Limitations of the study

Because of a series of constraints on the present study, including the uniform curriculum standard across the country (as noted in Chapter 1), limited instructional time, as well as the unchangeable schedule in the participating school, the present study may have been limited in the following aspects.

First and foremost, the research design may have been biased towards as well as biased against the Complementary Reading and the Free Reading groups.

(1) Possible biases towards the Complementary Reading group

The researcher was the class English teacher of the Complementary Reading group, which may have led to some bias towards the Complementary Reading group. Before the program, the Complementary Reading participants received comprehensive training in extensive reading and established positive reading

habits. Even after the reading habit was well established, the researcher spent more time with them, taking time to share extensive reading with them, such as interesting bits of what she had read from the graded readers. This would make them want to read more as they realized that knowledge gives power and that there is no time when one can say that he has read enough.

(2) Possible biases against the Complementary Reading group

The Matthew Effect was one of the biases against the Complementary Reading group. A survey done by Clay (1993) showed that when children at 9 years of age do not become competent readers, it would be very difficult for them to catch up. Furthermore, it also showed that when one is a poor reader in the first grade, one will most likely be a poor reader at the fourth grade. Unlike the Free Reading participants who volunteered to participate, the Complementary Reading participants entered this program by convenience sampling. Some of them may not be competent or willing readers at the beginning, which was possible to exert adverse effect on their extensive reading. This effect was modified by the pre-program training, the researcher's follow-up monitoring and their pleasant experience with extensive reading.

(3) Possible biases towards the Free Reading group

The bias against the Complementary Reading group, the Matthew Effect, was a bias towards the Free Reading group. It is important to realize that the possible higher reading proficiency and stronger motivation were likely to help the Free Reading participants in their academic endeavors, such as improving exam performances. When the beginning English language proficiency of the three groups had been strictly controlled for, the effect of motivation cannot be neglected. Numerous studies have shown that motivation greatly facilitates readers' engagement in reading (e.g., Guthrie & Wigfield, 2000). However, as described in Chapter 3, the participants in both the Complementary Reading group and the Free Reading group experienced motivation problems. They showed reluctance in different phases of the present study. Thus the stronger motivation of the Free Reading participants should be interpreted as their curiosity for something new and willingness to experience extensive reading at the beginning of the present study instead of consistent enthusiasm throughout

the program. The Free Reading group's stronger motivation contributed greatly to the establishment of the Free Reading experimental group as students cannot be forced to participate because of research ethics, but the stronger motivation did not necessarily continue throughout the program. The statistically significant difference detected in the pre-, mid-, post-tests and the EPER Placement Test B between the two experimental groups and the control group and the statistically significant difference between the two experimental groups could still be, to a large extent, attributed to extensive reading and the two different treatments.

(4) Possible biases against the Free Reading group

Student readiness was one of the biases against the Free Reading group. Although they volunteered to participate in the program, they came into the program having had practice with passages of only a few hundred words of length, because this was what was taught in their junior middle schools and what they had been taught to decode for the entrance examinations. It could be that pre-program training would have been ideal because extensive reading required them to adopt new strategies to cope with the volume of material that they were going to read.

The teacher arrangement was the second bias against the Free Reading group. As noted above, the researcher trained and shared a lot with the Complementary Reading participants, which was a luxury that the Free Reading participants generally did not have. In fact, the Free Reading participants in private complained that they did not receive their expected support and encouragement from their class English teachers.

In addition to the research design which had biases both towards and against the two experimental groups, testing was another limitation of the present study. The pre-, mid-, post- and delayed post-tests employed in the present study were developed by the local test bureau, rather than tests specifically intended for the present study. A direct consequence of this arrangement was that the test might have been fully sensitive to the gains from extensive reading. For example, in the vocabulary sections of the mid-, post- and delayed post-tests, only ten words were tested, which were far from enough to measure vocabulary gains through

extensive reading. For another example, in the reading comprehension sections in the pre-, mid-, post- and delayed post-tests, the participants' reading ability was measured by several short texts ranging from 100 to 300 words in length, which was not sufficient to examine whether the participants were fluent readers.

These two limitations, i.e., research design and testing, however, can hardly be fixed because of the curricular and time constraints noted above. First and foremost, the teaching arrangement of the sample school did not allow the researcher to establish three identical groups by random sampling just for her own research purpose. Equally, the testing system of the context in which the present study was implemented allowed for no possibility of administering tests intended specifically for extensive reading. The solution to this, in the future, is for extensive reading to claim its rightful place in the official syllabus and standard curriculum. Special teaching arrangements can then be made for extensive reading, and better tests intended for extensive reading can then be developed and administered.

The third limitation was that the extensive reading materials might not appeal to all the participants in the two experimental groups. Only a single set of graded readers was offered to each experimental group because of the limited research funding, and the books may not have adequately catered to the participants' diversified reading tastes. Similar to the above two limitations, more books can be made available if extensive reading gains official status, and more funding is provided for extensive reading. All of the three limitations noted above highlight the need for a central rather than a peripheral role of extensive reading in the English language syllabus and curriculum.

The fourth limitation of the present study was that this extensive reading program was implemented in Shanghai, the most developed city in mainland China. Its international outlook and greater openness to foreign cultures may have increased the chance that the students in the control group are exposed to a fair amount of English, which may have influenced their English language proficiency and behavior and attitude towards English language learning and reading. Nevertheless it is noteworthy that significant differences were found

between the extensive reading groups and the control group within the same cohort of Shanghai students.

Despite the limitations, the present study has demonstrated that extensive reading worked with Chinese students. The teaching situation for the present study was difficult. In such an EFL context, the students did not have the experience or habit of reading in English, nor did they have a wealth of knowledge in English to support their reading. However, the present study showed that extensive reading produced remarkable results in this challenging environment, which served as solid evidence in favor of extensive reading in this context. In addition, the curriculum developer and the school administration management do not need to worry about making radical changes to the classroom pedagogy in order to implement extensive reading, as the group that did the best was the group that underwent the least intervention in terms of pedagogic methodology.

5.3 Recommendations and further areas of inquiry

As discussed in Chapter 2, there has been little agreement on the definitions of extensive reading in the literature. Commonly cited are the top ten principles of extensive reading proposed by Day and Bamford's as presented below (Day & Bamford, 1998/2002).

The primary goal of this section is to discuss the extent to which the present study abided by these ten principles. A comparison of the major features of the two experimental treatments in the present study with the top ten principles of extensive reading (Day & Bamford, 1998/2002) will reveal similarities and differences between the two, which will lead to a better understanding of extensive reading in this context. The nature of this comparison is a retrospective evaluation of the present study, with an aim to make recommendations for future extensive reading projects in secondary schools in mainland China and to also suggest further areas of inquiries in this domain.

1. The reading material is easy.

The graded readers in the present study are categorized into seven levels of difficulty, ranging from 250 to 2,500 headwords. The seven levels of difficulty

are intended to cater for the individual EPER levels of the participants as measured by the EPER Placement test. The lexical gaps between each two levels range from 150 to 700 headwords. The lexical gap between Level 1 and Level 2 is 300 headwords, between Level 2 and Level 3 300 headwords, between Level 3 and Level 4 400 headwords, between Level 4 and Level 5 400 headwords and between Level 5 and Level 6 700 headwords. In this situation, a potential problem is that the transfer between each two levels may not be as smooth as expected. The recommendation is that more readers from a wider range of series are needed to build a more gradual and smooth transfer between each two levels, for example, a 50 headword gap between each two levels. This would entail an extensive reading program of a longer duration, providing students more time and a higher success rate. This in turn could increase motivation and further yield more positive results.

2. A variety of materials on a wide range of topics must be available.

The graded readers used in the present study comprise of nine genres. The nine genres have covered almost all genres available in the market. The participants made no complaint about the genres but they expressed preference for topics that are more closely related to their life and culture. The recommendation is that more titles, especially titles that are adapted for the Chinese context should also be included.

3. Learners choose what they want to read.

In this study, the Complementary Reading participants had to read what was assigned to them level by level, and they did not have the freedom to stop reading what failed to interest them. The Free Reading participants chose what they wanted to read from the 235 titles available. Because the Complementary Reading participants need to discuss their reading in class, they had to read the same reading materials; otherwise the in-class discussion could not be carried out if every participant read different titles. The recommendation is that titles that interest the majority should be promoted. They can be used as the class readers and serve as the basis for in-class discussion. Mini-group discussions in which group member that have read the titles provided a summary for their other groupmates can also be considered, because it serves as a rotational basis which

can help group members to improve their speaking skills. This mini-group discussion can then be followed by a more general discussion involving the entire class, with more global issues related to the readings being addressed.

Meanwhile, recommendations from the peers and teachers play a vital role in helping the readers locate appropriate materials; otherwise the readers are likely to miss some titles they feel interested in. Recommendations from others are not necessarily interpreted as deprivation of readers' freedom to make their own choices; on the contrary, recommendations from various parties are equally important, if not more important than, readers' self-choices.

4. Learners read as much as possible.

The Complementary Reading participants read around 5,000 words per day in a limited time because of a heavy academic workload. The Free Reading participants read as much as they wished. The participants who read more did not achieve better gains through extensive reading. A further area of inquiry in subsequent extensive reading programs should be reading quality.

5. The purpose of reading is usually related to pleasure, information and general understanding.

The reading materials were not used for detailed linguistic study. The Complementary Reading participants were unable to choose what they wanted to read and they were requested to finish reading 5,000 words per day and then write up a brief book report. It is assumed that the Complementary Reading participants read for information and general understanding. The Free Reading participants were able to locate what interested them and read as much as they wished. It is assumed that the Free Reading participants not only read for information and general understanding but also for pleasure. An area for further inquiry is whether reading for pleasure will result in better gains or not. Our null hypothesis is an affirmative "yes".

6. Reading is its own reward.

The essence of "*Reading is its own reward*" is that there should be few or no follow-up activities after reading. A brief summary of what they had read was the

follow-up activity for both the Complementary Reading group and the Free Reading group. Given that it is not possible to ask every participant to report what they read in class orally, a brief report is to the interest of both readers and teachers. The recommendation is that diversified and less time-consuming activities, e.g., Moodle Reader Quiz, can be used to maintain the participants' interest in extensive reading.

7. Reading speed is usually faster than slower.

The Complementary Reading group demonstrated substantial gains in terms of reading speed, and the Free Reading group also showed considerable gains in terms of reading speed. Although the Free Reading group read at a slightly lower speed than the Complementary Reading group, they made greater gains in accuracy in the present study. A further area of inquiry is whether reading speed is a variable for building accuracy.

8. Reading is individual and silent.

The Complementary Reading participants and the Free Reading participants read on their own, but there was follow-up discussion among the Complementary Reading participants. Reading can be individual and silent, and definitely out of the classroom (also for the Complementary Reading group) because of the limited time. A further area of inquiry is whether group discussions and sharing have a facilitative effect on extensive reading. Another area for future inquiry is how a reading community can be run effectively to promote extensive reading.

9. Teachers orient and guide their students.

The researcher oriented and guided the Complementary Reading participants throughout the program. There were two orientations for the Free Reading participants before the present study began. Book reports form the dominant communication channel between the researcher and the Free Reading participants during the program. Individual talks were arranged on a need basis but the students made rather limited use of this option. Our recommendation is that a teacher's consistent guidance and support is crucial for students' development. It is an area for further inquiry whether presence and amount of teachers' orientation and guidance will lead to different results.

10. The teacher is a role model of a reader.

The researcher was unable to read with, nor share her reading with, the Free Reading participants because they were from different classes in the senior one grade. Given that the Free Reading treatment was administered among students from different classes, the researcher was prevented from demonstrating herself as a role model for extensive reading. Future research can further investigate whether the teachers' role as a model extensive reader has an impact on the students.

In addition to the retrospections and recommendations based on Day and Bamford's principles of extensive reading, we wish to draw the following areas of inquiries to the attention of extensive reading researchers and practitioners.⁸⁶

1. How can an extensive reading program influence individual learners who can be categorized into different levels of readiness for extensive reading?
2. What sort of challenges and changes do the learners encounter across the different phases of an extensive reading program, and across the different levels of the readers they read?
3. Is extensive reading related to the learners' development of reading skills such as predicting, scanning, inferencing, reading critically and recognizing functions/uses of different types of words or discourse markers? If yes, how and why?
4. How can the learners' writing improvement be effectively assessed in an extensive reading program, particularly in the Chinese context where the current test instrument commonly used in secondary schools requires students to write with a required word amount and thus may not reveal the most remarkable changes in the subjects' writing ability, as they may be able to write more fluently and are more expressive?
5. How can extensive reading and intensive reading be blended effectively?

To summarize, numerous studies in both L1 and L2 settings provide very strong support in favor of extensive reading. However, "these findings have not persuaded many L2 curriculum developers and programme administrators, as

⁸⁶ The author gratefully acknowledges the two external examiners for suggesting these additional areas of inquiries for future research.

evidenced by a lack of commitment to extensive reading in many L2 instructional settings” (Grabe & Stoller, 2011, p. 120). The present study examined the effects of extensive reading in a Chinese context, and it has shown that benefits generated from an extensive reading program integrated into standard curriculum. The teaching situation of the present study was challenging, spanning only one third of the whole senior high school period, in an EFL environment where the students lacked experience, habit and linguistic foundation for reading in English. In this challenging environment, extensive reading has nevertheless shown positive results. This provides solid evidence in favor of extensive reading.

Appendices

Appendix A The pre-test

金山区 2010 学年分班考试

英语试卷

(本卷满分 100 分, 考试时间 90 分钟)

I. Grammar (10%)

Directions: Beneath each of the following sentences there are four choices marked A, B, C and D. Choose the one answer that best completes the sentence.

- As _____ announced in today's papers, the Shanghai Export Commodities Fair is also open on Sundays.
A. being B. is C. to be D. been
- The organization had broken no rules, but _____ had it acted responsibly.
A. neither B. so C. either D. both
- _____ it over again and you will find how foolish you have been.
A. Think B. Thinking C. To think D. Thought
- _____ you say you have been considering, it must be left to Harold do decide.
A. But B. Since C. Though D. No matter how
- Shoes of this kind _____ at least one year.
A. wears B. wear C. is worn D. are worn
- They couldn't buy the dictionary, because _____ of them had _____ money on them.
A. all, no B. any, no C. none, any D. no one, any
- It is not that people suffer more stress today; it's just that they think they _____.
A. are B. can C. do D. have
- The traffic signs show clearly that people _____ faster than the speed limit.
A. do not drive B. will not drive C. may not drive D. might not drive
- The pair of trousers, which _____ carefully made, sold at a high price in that store.
A. it was B. being C. was D. having been
- The _____ look on his face suggested that he _____ the news.
A. disappointed...had expected B. disappointed...hadn't expected
C. disappointing...hadn't expected D. disappointed...hasn't expected

II. Cloze (30%)

Directions: For each blank in the following passage there are four words or phrases marked A, B, C and D. Fill in each blank with the word or phrase that best fits the context.

(A)

Job stress has been known to cause heart problems in people who already have cardiovascular disease. Now Finnish scientists have shown that 1 in healthy people the pressure of work can cause damage.

High blood pressure, lack of 2, smoking and being overweight 3 to heart disease—a 4 killer in many industrialized countries.

But Mika Kivimaki, of the Finnish Institute of Occupational Health, and his colleagues, who studied the 5 histories of 812 healthy Finnish men and women in a metal industry company over 25 years, said job stress also plays an important role.

Workers who had the highest job-related stress levels at the start of the study were more than twice as 6 to die of heart disease, 7 the study published in The British Medical Journal.

Work stress 8 too much work as well as a lack of satisfaction and feeling undervalued and 9.

Many people work long hours but if the effort is 10 the stress is minimized. Kivimaki said job pressure is damaging when being overworked is 11 with little or no control, unfair supervision and few career opportunities.

The British Heart Foundation said the results 12 earlier research showing that people in

jobs with low control, such as manual workers, could be at greater _13_ of heart disease than other employees.

“It is _14_ for people to try to minimize levels of stress at work and for employers to _15_ people to have more control at work and to be rewarded for their success,” the foundation said in a statement.

- | | | | |
|------------------|------------------|-----------------|----------------|
| 1. A. though | B. seldom | C. even | D. thereby |
| 2. A. exercise | B. patience | C. enthusiasm | D. interest |
| 3. A. add | B. propose | C. introduce | D. contribute |
| 4. A. leading | B. plain | C. moderate | D. heavy |
| 5. A. personal | B. professional | C. medical | D. family |
| 6. A. much | B. likely | C. equal | D. stable |
| 7. A. due to | B. owing to | C. according to | D. contrary to |
| 8. A. contains | B. brings | C. proceeds | D. involves |
| 9. A. isolated | B. unappreciated | C. disconcerted | D. stimulated |
| 10. A. rewarding | B. improved | C. increasing | D. neglected |
| 11. A. sponsored | B. responded | C. threatened | D. combined |
| 12. A. indicate | B. simplify | C. support | D. overflow |
| 13. A. portion | B. chance | C. danger | D. risk |
| 14. A. advisable | B. predictable | C. profitable | D. comfortable |
| 15. A. persuade | B. transfer | C. allow | D. rescue |

(B)

Children who underachieve at school may just have a poor working memory rather than low intelligence, according to researchers who have _16_ the first tool to assess memory ability in the classroom.

The researchers from Durham University _17_ more than 3,000 primary school children of all ages and found that 10% of them -18- poor working memory, which seriously influences their learning. But the researchers found that teachers _19_ identify a poor working memory and often describe children with this _20_ as inattentive or less intelligent. Working memory is the _21_ to hold information in your head and handle it mentally--- for example adding up two numbers spoken to you by someone else _22_ using pen and paper or a calculator, or memorizing oral directions. Children at school need this _23_ on a daily basis for a variety of tasks, such as following teachers’ _24_ or remembering sentences they have been asked to write down.

The new _25_ --- a combination of a checklist and computer program--- will enable teachers to identify and _26_ children’s memory ability in the classroom from as early as four-year-old. Although the tools have already been tested successfully in 35 schools across the UK and have been translated into 10 languages, this is the first time they have been made widely _27_.

Teachers tend to _28_ children with poor working memories as having attention problems or “dreamers”, but the new test will allow them to screen children for the disorder. We are already beginning to see children in a different light, knowing more about the difficulties faced by children with weakened working memory. We realize that they are not _29_ , inattentive or underachieving, but children who simply need a different approach. We think these new ways of learning can help both the teacher and the children to _30_ complete their work.

- | | | | |
|-------------------|----------------|-----------------|----------------|
| 16. A. proposed | B. processed | C. produced | D. proceeded |
| 17. A. surveyed | B. lectured | C. conducted | D. inspected |
| 18. A. think of | B. care about | C. refer to | D. suffer from |
| 19. A. fairly | B. simply | C. rarely | D. only |
| 20. A. hobby | B. problem | C. imagination | D. activity |
| 21. A. ability | B. possibility | C. personality | D. quality |
| 22. A. with | B. through | C. by | D. without |
| 23. A. technology | B. memory | C. experience | D. exercise |
| 24. A. steps | B. permission | C. instructions | D. personality |
| 25. A. tool | B. conclusion | C. calculator | D. research |
| 26. A. ignore | B. assist | C. manage | D. assess |

- | | | | |
|---------------------|--------------|-----------------|----------------|
| 27. A. available | B. opposed | C. suitable | D. mobile |
| 28. A. satisfy | B. classify | C. identify | D. justify |
| 29. A. creators | B. thinkers | C. strangers | D. daydreamers |
| 30. A. successfully | B. carefully | C. thoughtfully | D. usefully |

III. Reading Comprehension (30%)

Directions: Read the following passages. Each passage is followed by several questions or unfinished statements. For each of them there are four choices marked A, B, C and D. Choose the one that fits best according to the information given in the passage you have just read.

(A)

Getting paid to talk about the World Cup is a great job. I'm not a football commentator though—just an English teacher in Japan.

I came to Japan two years ago, and didn't think I would stay, but Japan has that effect on you. People often end up living here longer than they planned. I think it's best to teach in a bigger city where there are other foreigners to mix with, rather than a small town where English teachers often complain of feeling like a goldfish in a bowl. Many people choose to live in Tokyo, of course, which is good for the nightlife factor. But I'd say that for general quality of living, cities of neither too larger nor too small, like sapporo where I live, are better choices.

I teach English privately, which means I'm my own boss. If you want to devote yourself to private teaching, it's well worth doing a TEFL course first, because your lessons will be much better for it. The problem with private teaching is finding students; it took me a year to build up a full schedule of private lessons, so I started out teaching in schools part-time. Most of my foreign friends here work full-time for big English conversation schools. The salary is fine to live on, but whether you can save money depends on how much going out and traveling you do here.

The schools are reluctant to take time off—even teachers with tickets for the England—Argentina game had trouble getting the day off.

- From the passage we know in Japan the writer likes to live in _____.
 A. a small town
 B. a big city
 C. a city of middle size
 D. Tokyo
- According to the writer, one had better _____ first to do private teaching.
 A. take a TEFL course
 B. find students
 C. build up a full schedule
 D. decide his or her own lessons
- The underlined sentence in the second paragraph implies that _____.
 A. Japan is good for nightlife
 B. They can teach English privately in Japan
 C. There are many foreigners in Japan
 D. Japan has something more active than expected

(B)

Ralph and Ilene hadn't been to a baseball game in about five years. They lived only 15 miles from the stadium, but the heavy traffic on game day made those 15 miles seem more like 60 miles. It took them about an hour to get to the stadium. Then, when the game was over, it took half an hour just to get out of the parking lot. Then the drive home was another hour. In other words, the traveling took longer than the game itself.

One day, they decided to go to a game. They drove south on the 110 freeway. The 110 is California's original freeway, full of twists and turns. Accidents occur daily. California drivers think yellow lights and sharp curves mean the same thing—speed up. The traffic was lighter than they expected. They arrived at the stadium 40 minutes before game time. They paid the \$8 parking fee, parked and locked the car, and walked to the main entrance.

Several individuals were standing around outside the stadium, looking casual but actually selling tickets secretly. They walked over to a man in a red cap.

Ralph's instincts were correct. The man had tickets for good seats at a fair price. Ralph gave the man \$45 and thanked him.

“Don’t thank me, my friend. Put your hands behind your back, please. You’re under arrest.”

“What?” Ralph was astonished. “What’s going on?”

“Buying scalped ticket is illegal in Los Angeles, and we’ve got a new mayor who wants us to enforce all the laws that bring in money.” said the man. “Don’t worry. We’ll have you back here right after we book you. You can pay the \$150 fine with your credit card. Have a nice day. Oh, and enjoy the game!”

4. Why didn’t Ralph and Ilene go to the baseball game very often?
 - A. Because they were not very interested in baseball.
 - B. Because they lived far away from the stadium.
 - C. Because it was time-consuming to go there.
 - D. Because it was often hard to find a place to park.
5. Why did many accidents occur on the 110 freeway according to the passage?
 - A. Because the traffic lights were broken.
 - B. Because traffic was heavy there.
 - C. Because the road conditions were poor.
 - D. Because people always drove too fast.
6. By saying “Ralph’s instincts were correct,” the author means that _____.
 - A. Ralph was lucky enough to get two tickets right before the game started.
 - B. Ralph was wise enough to choose the best two tickets from the man.
 - C. Ralph was good at bargaining so the he got the tickets at a fair price.
 - D. Ralph identified the right person who offered tickets at reasonable prices.
7. What’s the man in a red cap?
 - A. A ticket dealer
 - B. A police officer
 - C. A robber
 - D. Mayor of the city

(C)

More than half of teachers in a UK survey said they thought plagiarism from the Internet is a problem. Some students, they said, are so lazy they don’t even bother to take the ads off the cut-and-pasted text.

Fifty-eight percent of the teachers interviewed in the Association of Teachers and Lecturers (ATL) questionnaire had come across plagiarism among their pupils. Gill Bullen from Itchen College in Southampton for example said pieces handed in by two students were identical and “significantly better than either of them could have done.”

“Not only that, the essays given in didn’t quite answer the title question I had set.” A teacher from Leeds said: “I had one piece of work so obviously ‘cut and pasted’ that it still contained adverts from the Web page.”

Connie Robinson from Stockton Riverside College, Stockton, said: “With less able students it is easy to spot plagiarism as the writing style changes mid-assignment, but with more able students it is sometimes necessary for tutors to carry out Internet research to identify the source of the plagiarism.”

Mary Bousted, general secretary of the ATL, said: “Teachers are struggling under a mountain of cut-and-pasting to spot whether work was the student’s own or plagiarism.” She called for robust policies to fight plagiarism, and asked for help from exam boards and the government in providing resources and techniques to detect cheats.

But there was another side. “I have found once students clearly understand what plagiarism is, its consequences and how to reference correctly so they can draw on published works, plagiarism becomes less of a problem,” said Diana Baker from Emmanuel College, Durham. “I think the majority of students who engage in plagiarism do it more out of ignorance than the desire to cheat. They really want to succeed on their own merit.”

8. What does the word plagiarism probably mean?
 - A. Stealing other people’s property.
 - B. Cheating in school exams.
 - C. Copying others’ words or ideas.
 - D. Buying someone else’s essays.
9. Why is it easier to spot plagiarism with less able students?
 - A. They forget to take the ads off the cut-and-pasted text.
 - B. Their essays are much better than they could have done.
 - C. their mid-term essays don’t go with the title questions.
 - D. the language in their essays does not follow the same style.
10. What can we learn from Diana Baker’s words?
 - A. Plagiarism is not a serious problem for most schools in UK.

- B. Most students do not quite understand what plagiarism is.
 - C. Students should learn how to make good use of published works.
 - D. Students should be encouraged to achieve success on their own.
11. What is this passage mainly about?
- A. plagiarism from the Internet brings harm to university teachers and students.
 - B. Teachers should be provided with resources and techniques to fight plagiarism.
 - C. Teachers should help students to achieve success through hard work.
 - D. Plagiarism from the internet has become a pressing concern in UK schools.

(D)

When we can see well, we do not think about our eyes often. It is only when we cannot see perfectly that we come to see how important our eyes are.

People who are nearsighted can only see things that are very close to their eyes. Many people who do a lot of close work, such as writing and reading, become nearsighted. Then they have to wear glasses in order to see distant things clearly.

People who are farsighted face just the opposite problem. They can see things that are far away, but they have difficulty reading a book unless they hold it at arm's length. If they want to do much reading, they must get glasses too.

Other people do not see clearly because their eyes are not exactly the right shape. This, too, can be corrected by glasses. Some people's eyes become cloudy because of cataracts. Long ago these people often became blind. Now, however, it is possible to operate on the cataracts and remove them.

When night falls, colors become fainter to the eyes and finally disappear. After your eyes have got used to the dark, you can see better if you use the side of your eyes rather than the centers. Sometimes, after dark, you see a small thing moving to one side of you, which seems to disappear if you turn your head in its direction. This is because when you turn your head, you are looking at the thing too directly. Men on guard duty sometimes think they see something moving to one side of them. When they turn to look straight at it, they can not see it any more, and they believe they were mistaken. However, this mistake happens because the center of the eye, which is very sensitive in daylight, is not as sensitive as the sides of the eye after dark.

(Note: Answer the questions or complete the statements in No More Than Ten Words)

12. We don't know that our eyes are of great importance until _____.
13. According to the passage, who is most likely to be nearsighted, a tailor, a guard or a painter?
14. How can a person who is farsighted do much reading?
15. To see a small thing at night, it is better to look _____.

IV. Translation: (10%)

Directions: Translate the following sentences into English, using the words or phrases given in the brackets.

- 1 保护环境是每一个中国公民的职责。(It ...)
- 2 我第一次写作文时, 犯了许多语法错误。(make)
- 3 他一点也不知道那家企业为何被一家外国公司接管了。(take over)

V. Guided Writing: (20%)

Directions: Write an English composition in about 120 words according to the instructions given below.

小区门口贴着一张广告, 征求一位可以辅导初中数学的小老师。你决定写一封求职信。

Appendix B The mid-test

金山区 2010 学年第一学期期末考试
高一英语试卷
(本卷满分 100 分, 考试时间 90 分钟)
第 I 卷 (共 70 分)

I. Listening Comprehension: (20%)

Section A Short Conversations

Directions: In Section A, you will hear ten short conversations between two speakers. At the end of each conversation, a third voice will ask a question about what was said. The conversations and the questions will be spoken only once. After you hear a conversation and the question about it, read the four possible answers in your paper, and decide which one is the best answer to the question you have heard.

1. A. Operator. B. Salesman.
C. Post-office clerk. D. Accountant.
2. A. 9:13. B. 8:05. C. 9:30. D. 9:05.
3. A. He danced in his sitting room.
B. He turned around in his sitting room.
C. His sitting-room is quite big.
D. His sitting-room is quite small.
4. A. The ground is covered with snow now.
B. It is snowing heavily.
C. She has lost her shoes.
D. She does not like the shoes.
5. A. At home. B. In a coffee shop. C. On a street. D. In a car.
6. A. He went to an exhibition. B. He was ill.
C. He did not like fishing. D. He had to work.
7. A. Prepare some more vegetables. B. Take some more vegetables.
C. Help to eat more vegetables. D. Stop eating more vegetables.
8. A. Buying a new TV set.
B. Finding a new place to park the old furniture.
C. Buying new furniture.
D. Finding a new TV set.
9. A. Buy some football game tickets. B. Stay at home and watch TV.
C. Go to the football game. D. Stay at home and do some reading.
10. A. She is trying to find herself a new job.
B. She is too busy to go to the party.
C. She has received the invitation to the party.
D. She lives far away from the speakers.

Section B Passages

Directions: In section B, you will hear two short passages, and you will be asked three questions on each of the passages. The passages will be read twice, but the questions will be spoken only once. When you hear a question, read the four possible answers on your paper and decide which one would be the best answer to the question you have heard.

Questions 11 through 13 are based on the following passage.

11. A. When the fish get large enough.
B. When they grow the fish from eggs.
C. When the fish have been caught.
D. When they feed the fish.
12. A. To get the food to eat B. To enjoy the music.
C. To see who the owner is. D. To greet the owner.
13. A. To make an experiment. B. To produce more seafood.
C. To make a study of salt water fish. D. To make use of music.

Questions 14 through 16 are based on the following passage.

14. A. To get the answer sheet ready. B. To be sure you understand all the directions.

- C. To sharpen your pencil.
 D. To enter your name.
15. A. When you answer a specific question.
 B. When you take a final examination.
 C. When a test is scored by a computer.
 D. When a pen is not available.
16. A. Neatly and completely.
 B. Quickly but correctly.
 C. Slightly and carefully.
 D. Quickly but carefully.

Section C Longer Conversation

Directions: In section C, you will hear one longer conversation. The conversation will be read twice. After you hear the conversation, you are required to fill in the numbered blanks with the information you hear. Write your answers on your answer sheet.

Blanks 17 through 20 are based on the following conversation.

Complete the form. Write **ONE WORD** for each answer.

| | |
|---|--|
| Sender: Helen Lee Flat 5B, Greenwood Garden Shanghai, China Tel: (17) _____ | Receiver: Shirley Lee, 2-19 Suzuhari Cho Hadano City Tokyo, (18) _____ Tel: 045-743-8743 |
| Agreement No.: 7609 Customs Declaration: Contents: a (19) _____ Value: \$800 Net weight: 1 kg | Total Postage: (20) \$ _____ |

II. Grammar & Vocabulary (29%)

Section A

Directions: Beneath each of the following sentences there are four choices marked A, B, C and D. Choose the one answer that best completes the sentence.

21. Professors have a theory—8 a.m. classes are bad because ____ the teacher ____ the students are awake yet.
 A. either...or
 B. not only...but also
 C. both...and
 D. neither...nor
22. A great number of small power stations _____ up in this province since I started to work there.
 A. are set
 B. have set
 C. have been set
 D. will have been set
23. Go and get your books. They are _____ you left them.
 A. there
 B. there where
 C. where
 D. where there
24. He told me the problem might be the way _____.
 A. I communicate
 B. in that I communicate
 C. to communicating
 D. I was communicated
25. He was so shy that he _____ strangers the way.
 A. not dared ask
 B. dared not to ask
 C. dared not ask
 D. not dared to ask
26. Read English very often, _____ improve your English step by step.
 A. you will
 B. and you
 C. and you will
 D. and
27. Small sailboats can easily turn over in the water _____ they are not managed carefully.
 A. though
 B. before
 C. until
 D. if
28. This is a table that tells us what food experts advise us _____ to stay _____.
 A. eaten, healthy
 B. eating, health
 C. to eat, healthy
 D. eat, health
29. I do not think Jack saw me. He _____ into space.
 A. just stared
 B. was just staring
 C. has just stared
 D. had just stared
30. To keep safe in an earthquake, you should remain at a place ____ nothing may fall on you.
 A. whose
 B. which
 C. where
 D. that

31. When you are shopping online, all you have to do is to move the “mouse” from one Home Page to _____ and select something interesting to read.
A. other B. others C. another D. the others
32. Mary _____ come with use tonight, but she isn’t very sure yet.
A. must B. may C. can D. shall
33. Doing morning exercises _____ will do much good to your health.
A. ordinarily B. especially C. regularly D. generally
34. _____ along with host families, students are likely to get enough language practice.
A. Staying B. Having stayed
C. Stay D. To stay
35. She thought I was talking about her daughter, _____, in fact, I was talking about my daughter.
A. whom B. where C. which D. while
36. The doctor’s judgments proved _____ as the patient’s health conditions developed.
A. to be corrected B. to have been corrected
C. to be correct D. to correct
37. In the west, people make _____ a rule to send Christmas present to their relatives and friends.
A. there B. this C. that D. it
38. Mother is much kinder to the youngest child than to the others, _____ makes the other jealous.
A. than B. who C. of whom D. which
39. People read newspaper stories for a reason: to learn more about _____ they are interested in.
A. how B. that C. what D. which
40. Is it doing harm to personal freedom to insist that old people go to live with their relative _____ they might be taken better care of?
A. in case B. so that C. since D. until

Section B

Directions: Complete the following passage by using the words in the box. Each word can only be used once. Note that there is one word more than you need.

| | | | | |
|------------|------------|---------------|---------------|--------------|
| A. alone | B. mainly | C. artificial | D. comparison | E. contain |
| F. amusing | G. variety | H. raise | I. harmful | J. announced |

There are more than 37,000 known species of spiders in the world. They are in a (n) (41) of shapes and sizes. The largest spiders in the world live in the rain forest of South America known as the “bird-eating” spiders. They can grow up to 28 centimeters in length. In (42), the smallest ones are less than half a millimeter long and live in plants that grow on mountain rocks

Some people like to keep spiders as pets, particularly tarantulas which can live for up to 25 years. Most people, on the other hand don’t like touching spiders. Many people are afraid of spiders (43) because of their poison. In fact, only thirty of the 37,000 known species of spiders are dangerous to people. Spiders actually provide benefits to people, by catching and eating (44) insects such as flies and mosquitoes.

The main thing that makes spiders different from other animals is that they spin webs to catch the small insects they fell on. Some people have tried to (45) spiders commercially to collect the silk that these spiders produce, but no one has ever really managed to make a go of it. In addition, spiders are usually solitary (独居的) animals, and need to be kept (46).

One Canadian company may have found a solution to making (47) spider webs. In 2002, they (48) that they had used genetically modified (基因转换) goats to produce milk that may (49) the chemicals used to make spider webs.

III. Reading Comprehension (26%)

Section A

Directions: For each blank in the following passage there are four words or phrases marked A, B, C and D. Fill in each blank with the word or phrase that best fits the context.

In the past, man did not have to think about the protection of his environment. There

were few people on the earth, and natural resources seemed to be (50).

Today things are (51). The world has become too crowded. We are using up our natural resources too quickly, and at the same time we are (52) our environment with dangerous chemicals. If we continue to do this, human life on the earth will go to the end.

Many people (53) that if too many fish are taken from the sea, there will soon be none left. Yet, with modern fishing methods, more and more fish are caught. We know that if too many trees are cut down, (54) will disappear and nothing will grow on the land. Yet, we continue to use bigger and more (55) machines to cut down more and more trees.

If rivers are polluted with waste products from factories, we will die. (56), in most countries waste products are still put into rivers or into the sea, and there are few laws to stop this.

We know, too, that if the population of world continues to rise at the present rate, in a few years, there will not be enough (57). What can we do to solve these problems?

Scientists suggest that our natural resources will last longer if we learn to recycle them. And the world population will not rise so quickly if people use modern methods of birth (58). Finally, if we educate people to think about the problems, we shall have a better and cleaner (59) in the future.

- | | | | |
|-------------------|--------------|----------------|----------------|
| 50. A. beautiful | B. unlimited | C. rare | D. valuable |
| 51. A. common | B. similar | C. changeable | D. different |
| 52. A. protecting | B. saving | C. polluting | D. fighting |
| 53. A. wonder | B. realize | C. discover | D. invent |
| 54. A. mountains | B. seas | C. trees | D. forests |
| 55. A. powerful | B. dreadful | C. fashionable | D. convenient |
| 56. A. Thus | B. However | C. So | D. Furthermore |
| 57. A. people | B. animals | C. food | D. lives |
| 58. A. control | B. project | C. plan | D. reward |
| 59. A. natural | B. sea | C. planet | D. forest |

Section B

Directions: Read the following passages. Each passage is followed by several questions or unfinished statements. For each of them there are four choices marked A, B, C and D. Choose the one that fits best according to the information given in the passage you have just read.

(A)

Long before there wasn't any Washington or any United States, while there were trader Indians on the shores of the river that flows by the place where the present Washington D.C. stands. There Indians *paddled canoes* (划独木舟) up and down the river and traded with other Indians, exchanging things they had for things they wanted—furs for bread, bows for arrows, corn for potatoes. In the Indian language the same for traders was Potomac, so now we call the river the Potomac River after the trader Indians.

The Potomac separates two states – Maryland and Virginia. They are named after two English queens – Queen Mary and Queen Elizabeth I, who never married and was called the Virgin Queen. The Potomac Indians paddled their canoes down the river until they came to a much broader body of water. This body of water was so big that it seemed to them like the ocean, and they called it “the Mother of Waters”, which in their own language was “Chesapeake”. Though Chesapeake Bay is not the ocean, it is the biggest bay in the United States.

60. How did the Indians trade with each other?

- A. They paid money for what they wanted.
- B. They worked for others.
- C. They traded things they had for things they wanted.
- D. They gave things to others for free.

61. The articles states that Virginia and Maryland were named _____.

- A. after two Indian words
- B. after two states
- C. after two queens
- D. after two rivers

62. The word “Potomac” is Indian for _____.
 A. river
 B. the Mother of Waters
 C. Washington
 D. traders
63. According to the article, which of the following statements is correct?
 A. The Potomac Indian traders could not be trusted.
 B. The Potomac River was created by the Indians.
 C. The Chesapeake Bay looked like an ocean to the Indians.
 D. The Potomac Indian traders traded only with the people living in Washington.

(B)

(You may read the questions first.)



Best Weekend Projects

List Price: \$17.95

Our Price: \$13.96

You Save: \$3.99 (21%)

SKU

79460

In Stock

Usually ships in 1 business day

Choose from 80 unique ideas to create an extraordinary living space. The projects are practical, as well as attractive, and will **enhance** your home and yard, and can be made in a weekend.

In The Family Handyman's clear, dependable style, these 80 well-designed projects are presented in a clear, easy-to-follow style. Each project starts with an introduction to the tools, techniques, and materials, cost, and continues with detailed illustrated instruction plus tips and suggestions for making the projects easier and more efficient.

The projects are easy, intermediate, and advanced, indicating (说明) the degree of difficulty and skill level required, making it simple to choose the right project for your own ability. Throughout the book are sidebars—"Plan Smart," "Buy Smart," "Work Smart", cutting and materials lists, and professional tips from the editors. Plus many chapters include a collection of "Easy for Everyone" ideas.

Learn new tips and discover new ways to beautify and organize your home. Before you know it, you and your family will enjoy your new and improved home and yard after merely a weekend.....

64. How much does The Family handyman cost now?
 A.\$3.99
 B. \$13.96
 C.\$17.95
 D. \$21
65. We learn from the passage that the projects aim to _____.
 A. beautify and organize your home
 B. offer more jobs for the workless people
 C. sell tools and materials
 D. tell us how to take photographs
66. The word "**enhance**" in the passage most probably means "_____".
 A. decrease
 B. destroy
 C. improve
 D. support

(C)

Last night was the last game for my eight-year-old son's soccer team. It was the final quarter. The score was two to one, my son's team in the lead. Parents surrounded the playground, offering encouragement.

With less than ten second remaining, the ball suddenly rolled in front of my son's teammate, Mickey O'Donnel. With shouts of "Kick it!" echoing across the playground,

Mickey turned around and gave it everything he had. All around me the crowd erupted. O'Donnel had scored!

Then there was a silence. Mickey had scored all right, but in the wrong goal, ending the game in a tie. For a moment there was a total hush. You see, Mickey has Down syndrome (综合症) and for him there is no such thing as a wrong goal. All goals were celebrated by a joyous hug from Mickey. He had even been known to hug the opposing players when they scored.

The silence was finally broken when Mickey, his face filled with joy, hugged my son tightly and shouted, "I scored! I scored. Everybody won! Everybody won!" For a moment I held my breath, not sure how my son would react. I need not have worried. I watched, through tears, as my son threw up his hand in the classic high-five salute and started chanting, "Way to go Mickey! Way to go Mickey!" Within moments both teams surrounded Mickey, joining in the chant and congratulating him on his goal.

Later that night, when my daughter asked who had won, I smiled as I replied, "It was a tie. Everybody won."

67. What was the score of the score match?
A. Two to two, a tie to both teams.
B. Three to one in Mickey's team's favor.
C. Two to one in the opposite team's favor.
D. Everybody won because of Mickey's goal.
68. The underlined word "hush" in paragraph 3 means _____.
A. cheer B. cry C. laughter D. silence
69. What did the author worry about when Mickey scored and hugged his son?
A. The result of the match would fail his son.
B. His son would shout at Mickey for his goal.
C. Mickey would again hug the opposing player.
D. His son would understand Mickey's wrong goal.
70. It can be inferred from the passage that _____.
A. both teams liked and respected Mickey
B. both teams were thankful to Mickey for his goal
C. Mickey didn't mind though his goal was wrong
D. Mickey was a kind-hearted boy and hoped everybody won

Section C

Directions: Read the following text and choose the most suitable heading from A-F for each paragraph. There is one extra heading which you do not need.

- | |
|---|
| A. Advantages of Jet-ski B. Supervisor's Duty C. A Popular New Sport D. Injuries and Worse E. Rules are Produced F. Disadvantages of Jet-ski |
|---|

71.

Jet-ski is a kind of new sport. Its market is one of the fastest growing sections of the leisure industry. It becomes popular for many reasons. Jet-ski is exciting—people skim across the water at speeds of up to 100 kph. It is also easy to use and come in one, two or three-seats models.

72.

However, not everyone is happy about its popularity, largely because of the big noise Jet-ski creates and the accidents Jet-ski causes.

73.

In a large number of incidents around the world related to jet-ski, swimmers and other jet-skiers had been hurt or even killed by the machines. For instance, in 1999 two tourists visiting the Middle East were involved in a jet-ski accident in which one of them died.

74.

A middle-aged man from a small seaside town in New Zealand recently died of head injuries after his small boat was hit by a jet-ski. Following this tragedy, a group people organized and awareness campaign to warn people about the dangers of the machines. The group has produced a leaflet containing advice such as always wearing a life jacket, keeping away from other people in the water, never using a jet-ski at night, and learning how to operate your jet-ski correctly. These are simple rules, but it is surprising how few people follow them.

75.

The National Park Service of America (NPS) has decided to take more extreme action and jet-ski have been forbidden from some of the waterways in its parks. Individual park supervisors decide whether jet-skis pose a threat to visitors, wildlife and waterways. If they feel it is necessary, jet-skis will not be allowed. Therefore, park supervisors take the responsibility for the final decision regarding jet-skis.

第 II 卷 (共 25 分)

I. Translation: (15%) 注意: 本大题考生只需做一组

Directions: Translate the following sentences into English, using the words or phrases given in the brackets.

(普通中学考生做以下组)

1. 我花了一个小时写完这篇文章。(It)
2. 他喜欢农场生活, 而不喜欢城市生活。(prefer)
3. 他每天只吃蔬菜不吃肉。(instead of)
4. 他非常关注这个歌手的信息。(close attention)
5. 当她得知得了第一名时, 她欣喜若狂。(on top of the world)

(重点中学考生做以下组)

1. 他说喜欢农场生活, 而不喜欢城市生活。(prefer)
2. 我花了一个半小时来考虑如何写好这篇文章。(It...)
3. 当这名参赛者得知自己得了第一名时, 她欣喜若狂。(on top of the world)
4. 这栋楼起火时很多人并没有意识到危险。(aware)
5. 令我们惊讶的是他晚饭只吃苹果, 不吃肉和蔬菜。(instead of)

II. Guided Writing: (10%)

Directions: Write an English composition in about 100 words according to the instructions given below.

请以 “My Suggestion(s) On School Activities” 为题, 谈谈自己对学校学生活动的建议。你的作文必须包括:

- 你的建议
- 你的理由

(注: 文中不得出现所在学校的名称和人物的真实姓名)

Appendix C The post-test

金山区 2010 学年第二学期期末考试
高一英语试卷
(本卷满分 100 分, 考试时间 90 分钟)
第 I 卷 (共 75 分)

I. Listening Comprehension: (20%)

Section A Short Conversations

Directions: You will hear ten short conversations between two speakers. At the end of each conversation, a third voice will ask a question about what was said. The conversations and the questions will be spoken only once. After you hear a conversation and the question about it, read the four possible answers in your paper, and decide which one is the best answer to the question you have heard.

- | | |
|---|---|
| 1. A. In a library. | B. In a shopping center. |
| C. At the airport. | D. In the post office. |
| 2. A. 40. | B. 30. |
| C. 20. | D. 10. |
| 3. A. Take a long walk. | B. Take a good rest. |
| C. Read a love story. | D. Catch up with her work. |
| 4. A. He likes to listen to the radio. | |
| B. The present roommate is very kind to him. | |
| C. He needs a quiet place. | |
| D. The new apartment is cheaper. | |
| 5. A. A delivery boy. | B. A postman. |
| C. Their neighbor. | D. Their neighbor's daughter. |
| 6. A. On the table in the living room. | B. On the table in the bedroom. |
| C. Lily got it. | D. The woman got it. |
| 7. A. She repaired her television. | |
| B. She asked the store to repair her television. | |
| C. She asked the store to change her television with a new one. | |
| D. She helped the woman replace her television. | |
| 8. A. Because there was a car accident. | B. Because of the rain. |
| C. Because of the foggy weather. | D. Because of the delay of the flight. |
| 9. A. Child and parent. | B. Student and teacher. |
| C. Parent and teacher. | D. Secretary and boss. |
| 10. A. He found it enjoyable. | B. He found it difficult to understand. |
| C. He found it informative. | D. He found it boring. |

Section B Passages

Directions: You will hear two short passages, and you will be asked three questions on each of the passages. The passages will be read twice, but the questions will be spoken only once. When you hear a question, read the four possible answers on your paper and decide which one would be the best answer to the question you have heard.

Questions 11 through 13 are based on the following passage.

11. A. Because he didn't have enough exercise in shooting wild animals.
B. Because it was always raining.
C. Because his eyes were not very good and he had to wear glasses.
D. Both A and C.
12. A. A bear. B. A lawyer. C. A hunter. D. Himself.
13. A. Mr. Smith was invited to go shooting in the forests.
B. It was on a rainy day that Mr. Smith went shooting.
C. Mr. Grey is the lawyer.
D. The bear shot back at Mr. Smith.

Questions 14 through 16 are based on the following passage.

14. A. Teachers of the university.
B. People who have constructed the library.

- C. Students of the school.
D. Employees working in the library.
15. A. 14 years. B. 40 years.
C. Less than two years. D. Less than one year.
16. A. Sometime next month. B. Next Sunday.
C. Tomorrow evening. D. On December 24.

Section C Longer Conversation

Directions: You will hear one longer conversation. The conversation will be read twice. After you hear the conversation, you are required to fill in the numbered blanks with the information you hear. Write your answers on your answer sheet.

Blanks 17 through 20 are based on the following conversation.

Complete the form. Write **ONE WORD** for each answer.

| |
|---|
| David came to California on (17) _____. |
| He is leaving California next (18) _____. |
| Julie came to California on (19) _____. |
| They both stay at the same (20) _____. |

II. Grammar & Vocabulary (29%)

Section A

Directions: Beneath each of the following sentences there are four choices marked A, B, C and D. Choose the one answer that best completes the sentence.

21. _____ arriving in New York, I gave my mum a call to let her know I had a safe flight.
A. In B. At C. On D. With
22. Tom's maths is better than _____ in the class.
A. anyone's else B. anyone else C. anyone else's D. anyone's else's
23. This is Mr. Smith, _____ I think has something interesting to tell you.
A. who B. whom C. that D. /
24. _____ it is to have a walk after dinner!
A. How fun B. What a fun C. What funny D. What fun
25. I can't imagine _____ a girl of that sort.
A. to marry B. marrying C. married D. marry
26. I like the little girl very much, _____ she can be very annoying sometimes.
A. even though B. furthermore C. so that D. nevertheless
27. —Why didn't you attend the party?
—I'm sorry, but I _____ a parents' meeting at my daughter's school.
A. am having B. have C. have had D. was having
28. As the light turned green, I stood for a moment, not _____, and asked myself what I was going to do.
A. moved B. moving C. to move D. being moved
29. _____ is known to us all, every coin has two sides, so don't just focus on the disadvantages of this idea.
A. It B. What C. As D. Which
30. We are talking about _____ we will spend our summer vacation. Jim insists on going to Maldives, while Tony wants to go to Singapore.
A. what B. that C. when D. where
31. —Will he come to our party?
—I don't mind. It makes no difference _____ he will.
A. whether B. which C. what D. when
32. —Excuse me, Sir. But you are in my way.
—Oh, sorry, I _____ that at all.
A. haven't realized B. don't realize C. didn't realize D. hadn't realized
33. A girl presented her mother with a paper cup _____ she had drawn a heart and a kiss on Mother's Day.
A. for which B. on which C. which D. to which
34. Scientists are trying to develop a material _____ similar properties as spider silk.
A. having B. having had C. have D. to have

35. _____ you haven't done any harm to your friends, you don't have to have a guilty sense.
A. Whether B. Unless C. Since D. While
36. A library with five thousand books _____ to the nation as a gift.
A. is offered B. has offered C. are offered D. have offered
37. What do you think made Mary so happy?
_____ the first prize in the competition.
A. As she won B. Won
C. Winning D. Because of winning
38. Recently he bought an ancient Chinese vase, _____ was very reasonable.
A. which price B. the price of which
C. the price of whose B. the price of which
39. Please give the ticket to _____ you think is interested in the film.
A. who B. the one C. whoever D. whomever
40. I didn't know anything about it. Only when I telephoned her _____ me about the accident.
A. she did tell B. did she tell C. she told D. told she

Section B

Directions: Complete the following passage by using the words in the box. Each word can only be used once. Note that there is one word more than you need.

| | | | | |
|----------------|----------------|-------------|-------------|------------|
| A. find | B. traditional | C. disturbs | D. restore | E. useless |
| F. performance | G. carefully | H. hesitate | I. suddenly | J. shows |

Even if all *candidates* (应试者) arrived for their exams with equal knowledge and ability, some would manage to get higher grades because of technique. Sitting an exam is a skill in itself.

It is 41 to attempt to learn new material a day or two before examinations. After weeks of study it is important to clear the mind, not *cloud* (阻塞) it. Try not to change your diet or sleep patterns greatly. Research 42 that if students are used to having breakfast, their *academic* (学业的) 43 suffers when it is missed. And people who never eat breakfast do not do so well if one is 44 provided.

While it may be 45 to work with a cup of black coffee, be careful to avoid drinking too much caffeine during long hours of studying. It may harm rather than improve performance.

In the exam room, lay out your *equipment* (用具) carefully so you know where to 46 everything. Read right through the exam paper very 47. Check the instructions to ensure you are clear about exactly what you have to do.

Plan your time and decide how long you should spend on each question. If you attempt only three-quarters of the paper, you could lose a quarter of the *available* (可获得的) marks completely.

If you are unsure of anything, do not 48 to ask an *invigilator* (监考人). Complain if something 49 your concentration — someone tapping, a distant music lesson, or the invigilator's new shoes.

Avoid the temptation to stay with friends analyzing the paper in detail afterwards. Instead enjoy a spot of mental relaxation, give yourself a treat, and then move on to the next examination.

III. Reading Comprehension (21%)

Section A

Directions: For each blank in the following passage there are four words or phrases marked A, B, C and D. Fill in each blank with the word or phrase that best fits the context.

Different people have different views on sports and games. Some people seem to think that sports and games are 50 things that people do at times when they are not working. But in fact sports and games can be of great value, especially to people who work with their brains most of the day, and should not be treated only as amusements.

Sports and games build our bodies, prevent us from gaining weights, and keep us

healthy. But these are not their only ____ 51 _____. They give us valuable practice in helping eyes, brain and muscles work together. In table tennis, when our eyes see the ball coming, ____ 52 _____ its speed and direction, and pass this ____ 53 _____ on to the brain, the brain then has to decide what to do, and to send its orders to the muscles of the arms, legs and so on, ____ 54 _____ the ball is met and hit back where the ____ 55 _____ wants it to go. All this must happen with very great speed and only those who have had a lot of practice at table tennis can ____ 56 _____ this chain of events successfully. For those who work with their brains most of the day, the practice of such ____ 57 _____ is especially useful.

Sports and games are also very useful for ____ 58 _____. If each of the students learns to go all out for his team and not for himself on the sports field, he will later find it ____ 59 _____ to work for the good of society instead of only for his own benefit.

- | | | | |
|---------------------|-----------------------|---------------------|--------------------|
| 50.A. wonderful | B. unimportant | C. terrible | D. ideal |
| 51.A. ways | B. trends | C. uses | D. disadvantages |
| 52.A. slow | B. record | C. maintain | D. judge |
| 53.A. information | B. ball | C. event | D. progress |
| 54.A. now that | B. until | C. so that | D. unless |
| 55.A. order | B. muscle | C. trainer | D. player |
| 56. A. decide on | B. carry out | C. identify with | D. refer to |
| 57.A. skills | B. knowledge | C. language | D. methods |
| 58.A. body building | B. character training | C. skill developing | D. sports training |
| 59.A. unnecessary | B. intelligent | C. natural | D. amazing |

Section B

Directions: Read the following passages. Each passage is followed by several questions or unfinished statements. For each of them there are four choices marked A, B, C and D. Choose the one that fits best according to the information given in the passage you have just read.

(A)



Brigitta Schaeffler is a travel writer whose first language is German. She thought her English was excellent, until she began traveling. Now she's not so sure.

I had always prided myself on my learning of the language. After all, I had gone to university in England. None of them even had trouble understanding me, nor I them. But then I began to travel.

First, I travelled around England, where I only had to travel a short distance to find a totally different variety of English. Sometimes I couldn't understand a word a person was saying. I must say it was the older people I couldn't understand; most young people don't seem to use the local dialect. While having dinner in a little village in Lancashire, I *overheard* this conversation between two strangers in the restaurant (I've put the translation beside it):

| | |
|---|--|
| Woman: Eight up. Ast seen Billy? | <i>Hello. Have you seen Billy?</i> |
| Man: Tha wa? | <i>Pardon?</i> |
| Woman: My Billy | |
| Man: Oh Nay, but uz'll keep an eye out. | <i>Oh no, but we'll watch out for him.</i> |

When I lived in American, I was prepared for the difference in *accent* (口音) and vocabulary. I knew that sweets were candy, a lorry was a truck and that I mailed letters instead of posting them. I also knew I'd have to learn to spell differently, remembering to change words ending in "our" to "or" (e.g. colour – color) and s to z (e.g. recognise – recognize). But I hadn't realised there were also some grammatical differences.

60. *Brigitta Schaeffler* was proud of herself because _____.

- A. none of people in England could understand her
- B. she thought she was excellent at English
- C. she could find a totally different variety of English
- D. she had realized there were also some grammatical differences

61. *Brigitta Schaeffler* couldn't understand the older people because _____.

- A. they couldn't speak English clearly
- B. they appeared to use local dialect

- C. they preferred to use another language
D. there were some different spelling rules
62. The word **overheard** in the passage probably means ‘hear someone talking _____’.
A. patiently B. excitedly C. accidentally D. angrily
63. Which one is **NOT** mentioned in the text to show the differences in language itself?
A. Vocabulary B. Pronunciation C. Conversation D. Spelling

(B)

Percy and Florence Arrowsmith, who celebrated their 80th wedding anniversary on Wednesday, say the secrets of the world’s longest marriage are: don’t sleep on an argument, always share a kiss and hold hands before going to bed.

Percy Arrowsmith, 105, and his 100-year-old wife were married on June 1, 1925. They met at their church in Hereford in western England. Queen Elizabeth II sent her congratulations to the Arrowsmiths, who celebrated their anniversary at home with coffee and snacks and with family and friends. “**What a splendid achievement.** I send you my warm congratulations and best wishes for your 80th wedding anniversary,” the Queen’s card said.

The Arrowsmiths have three children, six grandchildren and nine great grandchildren. “He can’t settle down if I’m not holding his hand,” Florence Arrowsmith said, “I think we’re very blessed. We still love each other, that’s the most important part.” Asked for their secret, Florence said you must never be afraid to say “sorry”. The couple’s daughter Jane Woolley said her parents both very active. “She says she can’t dance any longer but she still can have a drink.”

Guinness World Records said the pair held record for the longest marriage for a living couple and the oldest total age of a married couple.

64. Florence Arrowsmith got married at the age of _____.
A. 25 B. 20 C. 18 D. 28
65. “What does “**a splendid achievement**” refer to?
A. Percy and Florence have 9 great grandchildren.
B. Percy and his wife still can drink.
C. Percy never had an argument with his wife.
D. Percy and Florence have been married for 80 years.
66. Which of the following is **NOT** the secret of the couple’s successful marriage?
A. Don’t go to sleep on an argument.
B. Always kiss each other before going to bed.
C. Always hold hands wherever they go.
D. Make an apology to the other when doing something wrong.
67. The passage mainly talks about _____.
A. the keys to the couple’s long marriage
B. the congratulation sent by Queen Elizabeth II
C. the happy family of the Arrowsmiths
D. an old couple’s celebrating 80th wedding anniversary

(C)

The cosmetics(化妆品) industry makes billions of dollars annually. Fashion and Hollywood have set the standards for “beauty” and most people agree with their taste. Women of all ages wear make-up! To most, it is important to be fashionable. But is this practice right or wrong? Could it even be sin(罪过)? Is painting your face “fashionable” to God?

Of course, some have no interest in what God says. The Bible has no meaning or value to them. They are not concerned with pleasing God. They are only interested in pleasing themselves or being accepted by people. Yet, others are concerned with what God says, but do not know **His** will.

The Bible is God’s Instruction Book to mankind. It reveals all the essential truths and principles necessary for salvation(拯救). However, most have rejected it as a source of authority in their lives. They are content to believe traditions without actually taking the time to prove why they do what they do. Most are also content to coast through life believing and

operating on assumptions. Others simply practice what they do because of habit—and old habits die hard.

Yet we must ask: Is the use of make-up something that only religious extremists(极端分子) worry about? Are “fanatics” the only ones who would consider such a question? Or is it something you should be concerned about?

You need to know what the Bible clearly states about this subject. If make-up is important to God, it has to be mentioned in His Word—both specifically and in principle. Wearing make-up is a worldwide custom—yet it is one that has become socially acceptable more recently than you may think. Therefore, to even question the use of make-up may seem old-fashioned or out-dated to countless millions of women—and, of course, men—who have grown accustomed to and comfortable with its widespread use.

68. What does the underlined word “His” in Paragraph 2 refer to?

- A. The God.
- B. The person wearing make-up.
- C. The Bible.
- D. Someone we don’t know.

69. The author, through the passage, intends to tell us _____.

- A. that make-up is widely used no matter whether God likes it or not
- B. the history of the development of cosmetics industry
- C. the relationship between religion and daily life
- D. film industry promoted cosmetics industry

70. What’s the author’s attitude towards wearing make-up?

- A. He’s in favor of it.
- B. He questions the use of it.
- C. His attitude is neutral(中立).
- D. He’s sick about it.

Section C

Directions: Read the following text and choose the most suitable heading from A-F for each paragraph. There is one extra heading which you do not need.

- A. Review regularly
- B. Make good use of class time
- C. Find a good place to study
- D. Plan your time carefully
- E. Scan before you read
- F. Take notes in class

Maybe you are an average student with an average intellect. You probably think you will never be a top student. This is not necessarily so. Anyone can become a better scholar if he or she wants to be. The following suggestions may be helpful to you.

71.

When you plan your work, you should make a list of things that you have to do. Be sure to set aside enough time to complete the work that you are normally assigned each week. Of course, studying shouldn’t occupy all of your free time. Don’t forget to set aside enough time for entertainment, hobbies and relaxation.

72.

Look around the house for a good study area. Keep this space, which may be a desk or simply a corner of your room, free of everything, but study materials. No radio or television! When you sit to work, concentrate on the subject!

73.

This means looking a passage over quickly but thoroughly before your begin to read it more carefully. Scanning a passage lets you preview the material and get a general idea of the content. Scanning will help you double your reading speed and improve your comprehension.

74.

Take advantage of class time to listen to everything the teacher says. Real listening in

class means less work later. Taking notes will help you remember what the teacher says.

75.

When you get home from school, go over your notes. Review the important points that your teacher mentioned in class. If you do it regularly, the material will become more meaningful, and you will remember it longer.

Develop a good attitude about tests. The purpose of a test is to show what you have learned about a subject. Tests do more than just provide grades; they let you know what you need to study more, and they help to make your new knowledge permanent.

第 II 卷 (共 25 分)

I. Translation: (15%)

Directions: *Translate the following sentences into English, using the words or phrases given in the brackets.*

1. 信心是成功的关键。(success)
2. 中日两国文化有很多相似之处。(common)
3. 他年纪太小了, 对如何存钱一无所知。(so...that...)
4. 昨天他弄丢了同桌的笔记, 很感内疚。(guilty)
5. 在班会课上, 老师经常强调诚实的重要性。(emphasize)

II. Guided Writing: (10%)

Directions: *Write an English composition in about 100 words according to the instructions given below.* (根据下列要求, 请写一篇 100 字左右的英语短文。注意不要出现真实的人名和校名)

A Special Gift of Mine

1. Describe your gift.(描述这个礼物)
2. Why is it so special? (为何如此特别?)

Appendix D Reading list

| Name | Level | Genre | Word count | Rating | No. of readers |
|--|-------|--------------------|------------|--------|----------------|
| Level 1 (400 Headwords) | | | | | |
| Sister Love and Other Crime Stories | 1 | Crime/Mystery | 5565 | 4.12 | 25 |
| Aladdin and the Enchanted Lamp | 1 | Fantasy/Horror | 5240 | 4.08 | 25 |
| A Ghost in Love and Other Plays Playscript | 1 | Playscripts | 5561 | 4.05 | 24 |
| A Little Princess | 1 | Human Interest | 5840 | 4.04 | 26 |
| Love or Money? | 1 | Crime/Mystery | 6010 | 4.04 | 25 |
| Sherlock Holmes and the Duke's Son | 1 | Crime/Mystery | 5800 | 4.04 | 25 |
| The Omega Files - Short Stories | 1 | Fantasy/Horror | 5830 | 3.99 | 27 |
| Little Lord Fauntleroy | 1 | Human Interest | 7250 | 3.96 | 22 |
| The Withered Arm | 1 | Classics | 5735 | 3.95 | 17 |
| The Wizard of Oz | 1 | Fantasy/Horror | 5440 | 3.94 | 14 |
| The Monkey's Paw | 1 | Fantasy/Horror | 4830 | 3.88 | 28 |
| White Death | 1 | Thriller/Adventure | 6600 | 3.84 | 17 |
| Titanic Factfile | 1 | Factfiles | 5529 | 3.83 | 18 |
| New York Factfile | 1 | Factfiles | 5140 | 3.76 | 30 |
| The Elephant Man | 1 | True Stories | 5400 | 3.76 | 27 |
| Goodbye Mr. Hollywood | 1 | Thriller/Adventure | 5200 | 3.75 | 24 |
| Christmas in Prague | 1 | Human Interest | 4720 | 3.72 | 23 |
| Deserts Factfile | 1 | Factfiles | 5519 | 3.72 | 24 |
| Sherlock Holmes and the Sport of Kings | 1 | Crime/Mystery | 5925 | 3.72 | 23 |
| The Adventures of Tom Sawyer | 1 | Classics | 5825 | 3.72 | 25 |
| The Lottery Winner | 1 | Human Interest | 5655 | 3.68 | 28 |
| The Phantom of the Opera | 1 | Fantasy/Horror | 6230 | 3.68 | 25 |
| Mary, Queen of Scots | 1 | True Stories | 6540 | 3.67 | 27 |
| Sherlock Holmes: Two Plays Playscript | 1 | Playscripts | / | 3.67 | 21 |
| Pocahontas | 1 | True Stories | 5320 | 3.66 | 26 |
| The Witches of Pendle | 1 | True Stories | 5730 | 3.66 | 17 |
| Scotland Factfile | 1 | Factfiles | 5150 | 3.65 | 22 |
| The Butler Did It and Other Plays Playscript | 1 | Playscripts | 5925 | 3.64 | 26 |
| The President's Murderer | 1 | Thriller/Adventure | 5270 | 3.60 | 22 |
| Animals in Danger Factfile | 1 | Factfiles | 5140 | 3.59 | 22 |
| England Factfile | 1 | Factfiles | 4640 | 3.56 | 24 |
| The Murder of Mary Jones Playscript | 1 | Playscripts | / | 3.56 | 25 |
| Five Short Plays Playscript | 1 | Playscripts | 4828 | 3.55 | 23 |
| London | 1 | Factfiles | 4800 | 3.55 | 29 |
| Remember Miranda | 1 | Human Interest | 5060 | 3.51 | 26 |
| Under the Moon | 1 | Fantasy/Horror | 5320 | 3.50 | 15 |
| One-Way Ticket - Short Stories | 1 | Human Interest | 5520 | 3.45 | 26 |
| Ned Kelly: A True Story | 1 | True Stories | 5775 | 3.43 | 30 |
| The Meaning of Gifts - Stories from Turkey | 1 | World Stories | 5254 | 3.42 | 26 |
| The Coldest Place on Earth | 1 | True Stories | 5500 | 3.34 | 26 |
| Mutiny on the Bounty | 1 | True Stories | 5825 | 3.33 | 29 |
| Level 2 (700 Headwords) | | | | | |
| Alice's Adventures in Wonderland | 2 | Classics | 6315 | 4.28 | 18 |
| Dead Man's Island | 2 | Thriller/Adventure | 5215 | 4.12 | 16 |
| Five Children and It | 2 | Fantasy/Horror | 5945 | 4.10 | 23 |
| The Murders in the Rue Morgue | 2 | Crime/Mystery | 6995 | 4.10 | 19 |
| Agatha Christie, Woman of Mystery | 2 | True Stories | 5955 | 4.09 | 16 |
| Death in the Freezer | 2 | Crime/Mystery | 6180 | 4.03 | 22 |
| Robinson Crusoe | 2 | Classics | 6830 | 3.93 | 24 |
| Voodoo Island | 2 | Fantasy/Horror | 5910 | 3.89 | 19 |
| Dracula | 2 | Fantasy/Horror | 7875 | 3.87 | 20 |
| Hamlet Playscript | 2 | Playscripts | 5000 | 3.86 | 16 |
| Much Ado About Nothing Playscript | 2 | Playscripts | | 3.85 | 17 |
| Ear-rings from Frankfurt | 2 | Thriller/Adventure | 6422 | 3.83 | 23 |
| The Beautiful Game Factfile | 2 | Factfiles | 7015 | 3.83 | 21 |
| Red Dog | 2 | True Stories | 7662 | 3.82 | 25 |
| The Canterville Ghost | 2 | Fantasy/Horror | 6100 | 3.82 | 26 |
| Anne of Green Gables | 2 | Human Interest | 5860 | 3.80 | 15 |
| Huckleberry Finn | 2 | Classics | 6180 | 3.77 | 26 |

| | | | | | |
|--|---|--------------------|-------|------|----|
| Love Among the Haystacks | 2 | Classics | 7030 | 3.77 | 26 |
| The Children of the New Forest | 2 | Human Interest | 6605 | 3.75 | 19 |
| Grace Darling | 2 | True Stories | 6685 | 3.72 | 24 |
| Seasons and Celebrations Factfile | 2 | Factfiles | 6508 | 3.68 | 20 |
| Marco Polo and The Silk Road | 2 | Factfiles | 6700 | 3.67 | 25 |
| Tales from Longpuddle | 2 | Classics | 6490 | 3.67 | 22 |
| Cries from the Heart - Stories from Around the World | 2 | World Stories | 6683 | 3.66 | 21 |
| Sherlock Holmes Short Stories | 2 | Crime/Mystery | 6280 | 3.66 | 22 |
| One Thousand Dollars and Other Plays Playscript | 2 | Playscripts | / | 3.64 | 24 |
| Romeo and Juliet Playscript | 2 | Playscripts | 6306 | 3.64 | 18 |
| The Death of Karen Silkwood | 2 | True Stories | 5585 | 3.63 | 16 |
| Changing their Skies - Stories from Africa | 2 | World Stories | 6987 | 3.60 | 20 |
| Henry VIII and his Six Wives | 2 | True Stories | 6310 | 3.60 | 23 |
| New Yorkers - Short Stories | 2 | Human Interest | 5895 | 3.57 | 28 |
| The Love of a King | 2 | True Stories | 6160 | 3.54 | 21 |
| Ireland Factfile | 2 | Factfiles | 7120 | 3.53 | 28 |
| The Year of Sharing | 2 | Thriller/Adventure | 6390 | 3.53 | 19 |
| Climate Change Factfile | 2 | Factfiles | 7151 | 3.52 | 19 |
| Rainforests Factfile | 2 | Factfiles | 6480 | 3.51 | 30 |
| Too Old to Rock and Roll and Other Stories | 2 | Human Interest | 5310 | 3.51 | 23 |
| A Stranger at Green Knowe | 2 | Human Interest | 6300 | 3.46 | 16 |
| The Jungle Book | 2 | Classics | 6510 | 3.45 | 20 |
| Stories from the Five Towns | 2 | Human Interest | 5540 | 3.44 | 23 |
| William Shakespeare | 2 | True Stories | 9135 | 3.41 | 16 |
| The Mystery of Allegra | 2 | Fantasy/Horror | 6115 | 3.38 | 20 |
| Return to Earth | 2 | Fantasy/Horror | 6250 | 3.35 | 26 |
| Matty Doolin | 2 | Human Interest | 6580 | 3.21 | 29 |
| The Importance of Being Earnest Playscript | 2 | Playscripts | / | 3.21 | 12 |
| The Piano | 2 | Human Interest | 6070 | 3.19 | 13 |
| The Pit and the Pendulum and Other Stories | 2 | Fantasy/Horror | 6000 | 3.05 | 11 |
| Level 3 (1,000 Headwords) | | | | | |
| Go, Lovely Rose and Other Stories | 3 | Human Interest | 8065 | 4.30 | 12 |
| Leaving No Footprint: Stories from Asia | 3 | World Stories | 11546 | 4.07 | 14 |
| Goldfish | 3 | Crime/Mystery | 10100 | 4.04 | 19 |
| Skyjack! | 3 | Thriller/Adventure | 8685 | 4.04 | 19 |
| Information Technology Factfile | 3 | Factfiles | 9614 | 4.02 | 15 |
| On the Edge | 3 | Thriller/Adventure | 10300 | 4.01 | 17 |
| Rabbit-Proof Fence | 3 | True Stories | 10600 | 4.01 | 21 |
| The Last Sherlock Holmes Story | 3 | Crime/Mystery | 9680 | 3.90 | 23 |
| Moondial | 3 | Fantasy/Horror | 10650 | 3.86 | 19 |
| Chemical Secret | 3 | Thriller/Adventure | 10150 | 3.85 | 21 |
| The Three Strangers and Other Stories | 3 | Classics | 11680 | 3.85 | 13 |
| The USA Factfile | 3 | Factfiles | 10188 | 3.83 | 11 |
| The Railway Children | 3 | Human Interest | 9295 | 3.82 | 15 |
| Love Story | 3 | Human Interest | 8755 | 3.79 | 20 |
| Recycling Factfile | 3 | Factfiles | 10098 | 3.78 | 18 |
| Tales of Mystery and Imagination | 3 | Fantasy/Horror | 11960 | 3.76 | 16 |
| Frankenstein | 3 | Fantasy/Horror | 9685 | 3.71 | 16 |
| Martin Luther King Factfile | 3 | Factfiles | 9871 | 3.69 | 17 |
| The Wind in the Willow | 3 | Classics | 11540 | 3.69 | 12 |
| The Picture of Dorian Gray | 3 | Fantasy/Horror | 10245 | 3.67 | 21 |
| As the Inspector Said and Other Stories | 3 | Crime/Mystery | 9600 | 3.66 | 9 |
| The Secret Garden | 3 | Human Interest | 10715 | 3.66 | 16 |
| The Star Zoo | 3 | Fantasy/Horror | 8915 | 3.65 | 17 |
| The Crown of Violet | 3 | Thriller/Adventure | 10800 | 3.62 | 18 |
| A Christmas Carol | 3 | Classics | 10385 | 3.59 | 17 |
| Ethan Frome | 3 | Classics | 10700 | 3.58 | 16 |
| Justice | 3 | Thriller/Adventure | 10420 | 3.55 | 19 |
| The Prisoner of Zenda | 3 | Thriller/Adventure | 10710 | 3.54 | 19 |
| The Everest Story | 3 | Factfiles | 10150 | 3.53 | 18 |
| The Call of the Wild | 3 | Classics | 10965 | 3.50 | 18 |

| | | | | | |
|---|---|--------------------|-------|------|----|
| Kidnapped | 3 | Thriller/Adventure | 12435 | 3.49 | 19 |
| The Brontë Story | 3 | True Stories | 10600 | 3.46 | 19 |
| Playing with Fire: Stories from the Pacific Rim | 3 | World Stories | 12589 | 3.44 | 17 |
| The Long White Cloud - Stories from New Zealand | 3 | World Stories | 11150 | 3.44 | 17 |
| A Pair of Ghostly Hands and Other Stories | 3 | Fantasy/Horror | 9600 | 3.42 | 12 |
| Through the Looking Glass | 3 | Classics | 10605 | 3.42 | 13 |
| Dancing with Strangers - Stories from Africa | 3 | World Stories | 11990 | 3.41 | 13 |
| The Card | 3 | Human Interest | 11100 | 3.40 | 20 |
| Wyatt's Hurricane | 3 | Thriller/Adventure | 10500 | 3.36 | 16 |
| Australia and New Zealand Factfile | 3 | Factfiles | 10270 | 3.34 | 14 |
| Who, Sir? Me, Sir? | 3 | Human Interest | 10295 | 3.24 | 12 |
| A Cup of Kindness: Stories from Scotland | 3 | World Stories | 11205 | 3.19 | 14 |
| Tooth and Claw - Short Stories | 3 | Human Interest | 8255 | 2.61 | 9 |
| Level 4 (1,400 Headwords) | | | | | |
| Death of an Englishman | 4 | Crime/Mystery | 14815 | 4.11 | 19 |
| The Eagle of the Ninth | 4 | Thriller/Adventure | 14950 | 3.98 | 16 |
| Washington Square | 4 | Classics | 15490 | 3.98 | 19 |
| A Morbid Taste for Bones | 4 | Crime/Mystery | 15300 | 3.94 | 21 |
| Black Beauty | 4 | Human Interest | 15400 | 3.93 | 21 |
| Disaster! Factfile | 4 | Factfiles | 14791 | 3.93 | 22 |
| Treasure Island | 4 | Thriller/Adventure | 15125 | 3.93 | 19 |
| The Big Sleep | 4 | Crime/Mystery | 15960 | 3.91 | 20 |
| The Thirty-Nine Steps | 4 | Thriller/Adventure | 17170 | 3.89 | 19 |
| Dr Jekyll and Mr. Hyde | 4 | Fantasy/Horror | 12520 | 3.88 | 26 |
| Three Men in a Boat | 4 | Human Interest | 18055 | 3.88 | 17 |
| Reflex | 4 | Thriller/Adventure | 16500 | 3.87 | 19 |
| The Hound of the Baskervilles | 4 | Crime/Mystery | 19330 | 3.85 | 23 |
| Persuasion | 4 | Classics | 19370 | 3.81 | 23 |
| Gulliver's Travels | 4 | Classics | 15325 | 3.79 | 20 |
| The Scarlet Letter | 4 | Classics | 15965 | 3.77 | 21 |
| We Didn't Mean to Go to Sea | 4 | Thriller/Adventure | 14860 | 3.77 | 19 |
| Cranford | 4 | Classics | 15015 | 3.76 | 14 |
| The Moonspinners | 4 | Thriller/Adventure | 14930 | 3.76 | 16 |
| Mr. Midshipman Hornblower | 4 | Thriller/Adventure | 14700 | 3.76 | 21 |
| Gandhi | 4 | Factfiles | 17000 | 3.75 | 24 |
| The Silver Sword | 4 | Thriller/Adventure | 14960 | 3.75 | 17 |
| Nelson Mandela Factfile | 4 | Factfiles | 16390 | 3.74 | 20 |
| Desert, Mountain, Sea | 4 | True Stories | 15580 | 3.73 | 22 |
| The Whispering Knights | 4 | Fantasy/Horror | 17100 | 3.73 | 18 |
| A Dubious Legacy | 4 | Human Interest | 15100 | 3.68 | 22 |
| Lord Jim | 4 | Classics | 19160 | 3.67 | 22 |
| The Unquiet Grave | 4 | Fantasy/Horror | 15860 | 3.67 | 17 |
| A Tale of Two Cities | 4 | Classics | 14850 | 3.65 | 25 |
| Lorna Doone | 4 | Human Interest | 17000 | 3.64 | 28 |
| Doors to a Wider Place - Stories from Australia | 4 | World Stories | 15862 | 3.63 | 23 |
| Little Women | 4 | Human Interest | 14920 | 3.63 | 22 |
| The African Queen | 4 | Human Interest | 15250 | 3.56 | 17 |
| The Songs of Distant Earth and Other Stories | 4 | Fantasy/Horror | 15210 | 3.55 | 14 |
| The History of the English Language Factfile | 4 | Factfiles | 16753 | 3.49 | 18 |
| Great Crimes Factfile | 4 | Factfiles | 15747 | 3.47 | 20 |
| Silas Marner | 4 | Classics | 16065 | 3.47 | 24 |
| Land of My Childhood - Stories from South Asia | 4 | World Stories | 16937 | 3.45 | 21 |
| The Price of Peace: Stories from Africa | 4 | World Stories | 16068 | 3.00 | 15 |
| Level 5 (1,800 Headwords) | | | | | |
| The Age Of Innocence | 5 | Classics | 24820 | 4.38 | 14 |
| Great Expectations | 5 | Classics | 24045 | 4.30 | 12 |
| Sense and Sensibility | 5 | Classics | 24345 | 4.26 | 11 |
| I, Robot | 5 | Fantasy/Horror | 22500 | 4.20 | 9 |

| | | | | | |
|---|---|--------------------|-------|------|----|
| Brat Farrar | 5 | Thriller/Adventure | 24510 | 4.08 | 11 |
| The Accidental Tourist | 5 | Human Interest | 24810 | 4.06 | 7 |
| Far From The Madding Crowd | 5 | Classics | 24490 | 4.03 | 9 |
| Wuthering Heights | 5 | Classics | 23180 | 4.00 | 10 |
| King's Ransom | 5 | Crime/Mystery | 24330 | 3.98 | 10 |
| Deadlock | 5 | Crime/Mystery | 22610 | 3.93 | 12 |
| Jeeves and Friends | 5 | Human Interest | 22670 | 3.84 | 7 |
| The Bride Price | 5 | Human Interest | 22620 | 3.83 | 16 |
| Ghost Stories | 5 | Fantasy/Horror | 22720 | 3.80 | 8 |
| Heat and Dust | 5 | Human Interest | 24125 | 3.70 | 6 |
| The Dead Of Jericho | 5 | Crime/Mystery | 27170 | 3.58 | 12 |
| David Copperfield | 5 | Classics | 24770 | 3.38 | 12 |
| The Riddle of the Sands | 5 | Thriller/Adventure | 22885 | 3.33 | 12 |
| The Garden Party and Other Stories | 5 | Human Interest | 22665 | 3.25 | 6 |
| This Rough Magic | 5 | Thriller/Adventure | 24750 | 2.73 | 11 |
| Do Androids Dream Of Electric Sheep? | 5 | Fantasy/Horror | 31300 | 0.00 | 0 |
| Treading on Dreams - Stories from Ireland | 5 | World Stories | 25395 | 0.00 | 0 |
| Level 6 (2,500 Headwords) | | | | | |
| A Passage to India | 6 | Human Interest | 29773 | 5.00 | 1 |
| Vanity Fair | 6 | Classics | 32940 | 5.00 | 2 |
| Jane Eyre | 6 | Classics | 31360 | 4.00 | 1 |
| The Joy Luck Club | 6 | Human Interest | 31120 | 4.00 | 1 |
| Oliver Twist | 6 | Classics | 26560 | 4.00 | 1 |
| Pride and Prejudice | 6 | Classics | 29455 | 4.00 | 1 |
| The Woman in White | 6 | Thriller/Adventure | 31770 | 4.00 | 2 |
| American Crime Stories | 6 | Crime/Mystery | 26500 | 0.00 | 0 |
| Barchester Towers | 6 | Classics | 29520 | 0.00 | 0 |
| Cold Comfort Farm | 6 | Human Interest | 32850 | 0.00 | 0 |
| Cry Freedom | 6 | True Stories | 29420 | 0.00 | 0 |
| Deadheads | 6 | Crime/Mystery | 26640 | 0.00 | 0 |
| Decline and Fall | 6 | Human Interest | 31245 | 0.00 | 0 |
| Dublin People | 6 | Human Interest | 26580 | 0.00 | 0 |
| The Enemy | 6 | Thriller/Adventure | 28850 | 0.00 | 0 |
| The Fly and Other Horror Stories | 6 | Fantasy/Horror | 28930 | 0.00 | 0 |
| Meteor | 6 | Fantasy/Horror | 26380 | 0.00 | 0 |
| Night Without End | 6 | Thriller/Adventure | 26720 | 0.00 | 0 |
| Tess of the d'Urbervilles | 6 | Classics | 33060 | 0.00 | 0 |

Appendix E Consent Form (English Version)

Consent Form

Dear students,

You are invited to participate in this research program entitled: Extensive Reading in Secondary Schools. Before you decide whether to participate or not, please read the following information as carefully as possible, which will help you understand this research, why this research is conducted, the procedures and duration of this research, the potential benefits and risks associated with participating in this research. If you are willing to, you may discuss it with your family and friends, or ask the research to make explanation to help you make the decision.

Extensive reading has been demonstrated to exert confounding and long-lasting effects on learners' listening, grammar, vocabulary, reading and writing in domestic and international research literature. The purpose of this research is to examine the effect of extensive reading on English language learning of senior high school students in Shanghai. You will be asked to do three things in this research: read the graded readers for this research under the guidance of the researcher, write book reports and complete the questionnaires. You may find your experience in this program enjoyable and maybe helpful to you improve your English language proficiency. The information you provided in this study may also help the researcher better understand the effect of extensive reading on English language learning of senior high school students in Shanghai

This research will be conducted in the two semesters in the 2010/11 academic year, namely the first semester from Sept 2010 to Jan 2011 and the second semester from Feb 2011 to Jun 2011. If you are not used to extensive reading, you may withdraw at any time. If your academic performance is negatively influenced by this research, the researcher will provide free make-up tutorial for you in the summer holiday in 2011.

I have read and understand the information on the form and I consent to volunteer to be a subject in this study. I understand that my responses are completely confidential and that I have the right to withdraw at any time. I have received an unsigned copy of this informed Consent Form to keep in my possession.

Name (Print)

Signature: _____

Date: _____

I certify that I have explained the above individual the nature and purpose, the potential benefits, and possible risks associated with participating in this research study, have answered any questions that have been raised, and have witnessed the above signature.

Signature: _____

Date: _____

Appendix F Consent Form (Chinese Version)

知情同意书

亲爱的同学：

您好！研究者邀请您参加此项高中泛读研究:Extensive Reading in Secondary Schools.在您决定是否参加这项研究之前，请尽可能仔细阅读以下内容。它可以帮助您了解该项研究以及为何要进行这项研究，研究的程序和期限，参加研究后可能给您带来的益处和风险。如果您愿意，您也可以和您的亲属、朋友一起讨论，或者请研究者给予解释，帮助您做出决定。

泛读(Extensive Reading) 已经在国内外的各项研究中被证明对于学习者的听力、语法、词汇、阅读、写作等方面有积极而深远的影响。此次在本校进行的研究，目的在于检测泛读对于上海高中生英语学习的影响。在这项研究中你会被要求完成以下三件事：你会在研究者的指导下阅读本项研究专门提供的英语分级读物，完成读书报告以及接受研究者的访问和问卷调查。您将会发现此次学习经历非常有趣，并且对于提高您的英语水平很有帮助。你在此项研究中所提供的信息将有助于我们更好的了解泛读对于上海高中生英语学习的影响。

此项研究将在 2010-2011 学年的两个学期进行，即 2010 年 9 月至 2011 年 1 月为第一学期，2011 年 2 月至 2011 年 6 月为第二学期。如您在研究过程中感到对泛读不适应，可随时退出。如果您因为不适应泛读而导致成绩下降，研究者愿意在 2011 年暑假为您提供免费的补习。

我已经阅读并理解了以上信息。我同意自愿成为这项研究中的一名被研究者。我明白我在这项研究中的所有信息都是完全保密的，并且我有权在任何时候退出这项研究。

姓名 _____

签名 _____

日期 _____

我声明我已经向上述个人解释了该项研究的实质，目的，潜在的好处以及参与此项研究可能造成的风险，并回答了所有被提出的问题。我见证了以上签名。

签名_____

日期 _____

Appendix G A preliminary discussion on graded readers

The first note to make at the beginning of this section is that the discussion on graded readers are not limited to the Oxford Bookworm series, which was read in the present study. The discussion will involve several major international series, such as Black Cat, Cambridge, Longman, Macmillan and Penguin, in order to present a cross-series dimension in discussion on graded readers. Some specific cases presented in this section will visualize the argument in Chapter 2, such as the authenticity of graded readers and the difference between simplification and rewriting. The second note is that the discussion and suggestions presented in this section are the observations and thoughts from the researcher herself. They are not views of any other parties. The third note is that some issues introduced and discussed in this section may not be called suggestions; instead, they are areas for further inquiry as solutions to these issues have not been found in the present study.

1 Issues involved in simplification

In the researcher's point of view, the quality of a graded reader should include at least two components: its language quality and story quality, i.e. readability. The importance of the language quality of the graded readers is understandable, as they are closely-related to language acquisition. The language quality of graded readers have been discussed in Chapter 2, with a conclusion that graded readers afford authentic language experience with an additional merit of being easy to process for learners of English. Our discussion in this section focuses on the readability of graded readers, namely how to present a more readable story within limited lexis and syntax.

Graded readers are texts written by means of lexical, structural, and information control (Hedge, 1985; Hill, 1997a/2008). Admittedly, 2,500 headwords allow more vivid description and thus make a better story than 1600 headwords. A comparison between two versions of *Pride and Prejudice* intuitively shows the difference. In this comparison, the *Pride and Prejudice* from OBW Level 6 and Macmillan Readers (MR) Level 5 are chosen. The OBW version is written with 2,500 headwords and the MR version 1,600 key words.

1. Opening sentence:

Original text: *It is a truth universally acknowledged that a single man in possession of a large fortune must be in want of a wife. However little known the feelings or views of such a man may be on his first entering a neighborhood, this truth is so well fixed in the minds of the surrounding families, that he is considered the rightful property of someone or other of their daughters.*

MR: Everyone knows that a rich, unmarried man needs a pretty wife. And every mother wants her daughters to be happily married.

OBW: It is a truth well known to all the world that an unmarried man in possession of a large fortune must be in need of a wife. And when such a man moves into a neighborhood, even if nothing is known about his feelings or opinions, this truth is so clear to the surrounding families, that they think immediately as the future husband of one or other of their daughters.

2. Mr. Collins' proposal to Charlotte Lucas

Original text: *But here she did injustice to the fire and independence of his character, for it led him to escape out of Longbourn House the next morning with admirable slyness, and hasten to Lucas Lodge to throw himself at her feet. In as short a time as Mr. Collins's long speeches would allow, everything was settled between them to the satisfaction of both.*

MR: The following morning, Mr. Collins left Longbourn House very early. He hurried to Lucas Lodge. Charlotte was looking out of an upstairs window. She left the house immediately and went to meet Mr. Collins. Mr. Collins asked her to marry him and Charlotte accepted at once.

OBW: He got up early the next morning and escaped from Longbourn House, in a great hurry to reach Lucas Lodge and throw himself at her feet. She did

not keep him waiting for an answer, and the happy couple found themselves engaged as quickly as Mr. Collins' long speeches would allow.

In this case, the OBW version is closer to the original in terms of writing style. The aim of this comparison is to show that more lexis and syntax allow more freedom in writing a graded reader, and in turn helps to reserve some important qualities of the original text, such as the author's humor. It is not the editors and writers to blame for the plain text. In some other cases, the loss of the original style or some important information is not caused by the lexis and syntax limit, as shown in the following example in *Silas Marner*. *Silas Marner* is published by Longman Fiction (LF) at Level 4 with 1000 headwords and OBW at Level 4 with 1400 headwords.

LF: Nancy did not want an adopted child.

OBW: Nancy's religion was extremely important to her. She firmly believed that people should accept whatever happened to them in life, because it was God who decided everything.

The OBW version makes it clearer to the readers why Nancy does not want to adopt a child. In the LF version, Nancy appears to be stiff and stubborn but in fact she is kind and loving to others. This comparison suggests that fewer headwords do not necessarily mean less information, simplified characters or lower readability. A few more words to maintain the important information can still present a good story within limited headwords.

Another typical example of the loss of original after simplification can be found in *Beowulf* published by Black Cat. *Beowulf* is the first European dialects epic, the oldest work written in old English and one of the most important works of Anglo-Saxon literature, which make it the subject of much scholarly study, theory, speculation, and discourse. The publication of *Beowulf* in a graded reader series is well worth praising because of the effort to expose readers to a variety of literature genres and introduce epics into graded texts. Poetry is a treasure trove of literature, and we cannot find any reason to exclude it from graded texts and

extensive reading.

As mentioned in foreword of *Beowulf* published by Black Cat, the use of alliteration is one of the discourse features of this work. An excerpt of simplification is presented below:

Original text

Now Beowulf bode in the burg of the Scyldings, leader beloved, and long he ruled in fame with all folk, since his father had gone away from the world, till awoke an heir.

Black Cat version

One of these kings was called Shield Sheafson.....Shield Sheafon's son was called Beow. Beow recognised what the people had suffered in the past and he rewarded them generously whilst his father was still king.

Although the simplification robs many of its original, it is successful in making the story clearer and more accessible to young readers. In this case, it is not simplification to blame. Without simplification, it is impossible for non-native speakers of English to fully understand and appreciate this great work without adaptation, let alone alliteration.

The third observation is made in *The Women in White*. In this comparison, the Chinese version of *The Women in White* is a faithful translation from the original. Its Chinese version published by Foreign Literature Press in 1983 is 745 pages, 488,000 Chinese words. The English version refers to *The Women in White* in OBW series at Level 6 written with 2,500 headwords, with a total word count of 31,770 English words.

Judging from the overall quality, the OBW *The Women in White* presents a very successful simplification. It makes a long story short by taking away some information, e.g., how Marian managed to rescue Laura from the asylum, but not at the expense of integrity and readability. For readers who have not read the original text, this story is complete as well as highly readable and fascinating. Its

success has been admitted by the participant in the present study. The complementary reading participant who read this book rated it as five, meaning it is an excellent title in this series.

On the premise that it is a successful simplification, there are still minor issues which may threaten the overall quality of the whole book. In terms of writing style, the original *The Woman in White* is considered to be among the first mystery novels using multiple narratives. This writing style has been kept in the OBW version, but quite disappointedly Count Fosco's narrative, which is supposed to be a climax of this story, is removed. It is indeed puzzling why the editors do not stick with the original writing style throughout the whole book since they seem to have decided to keep the multiple narratives.

In terms of individual characters, for example, Marian is described as one "of the finest creations in all Victorian fiction" by John Sutherland. Even the Count greatly admires Marian for her intellect in his narrative. In fact, the Count and Marian are the two most important features of this novel, much more impressive and successfully portrayed than Laura, pure and innocent but delicate, and Sir Percival, cold and wicked but superficial. To a large extent this story is a trial between Marian and Count Fosco. Unfortunately, since Count Fosco's narrative and part of Marian's effort is omitted, their roles are not as important as they originally are. The story is distorted in this sense. For another example, Professor Pesca, although not the leading man, he does not play a walk-on part in this story. He finds Hartright the Limmeridge job, introduces him to Laura and Marian, and proves Fosco's unexpected enemy. The original description of him as a lovely Italian is much more interesting and vivid than that in this book. But due to understandable reasons he is less lovely and interesting in the OBW version.

The third problem results from the classification of *The Women in White* in the OBW series. As noted above, *The Women in White* is categorized as a mystery novel, but in the OBW series it is categorized as a *Thriller & Adventure*. This is particularly puzzling when OBW has a *Crime & Mystery* catalogue. There must be reasons for the editors to do so, but on the surface there is difference between the original version and the OBW version.

Our stance in this discussion section is not to criticize simplification. To present a satisfactory simplification, effort had been made, but the result was not satisfactory because of the high complexity involved in this task.

A desire to be faithful to the original led early simplifiers just to cut words and phrases and sentences out. They just used scissors or a black pen. Great chunks were left as they were. The result was pretty awful.

The rewriters sought to preserve the meaning and style of the original but produce a soup of their own that above all made sense and held together in its own terms. To do this well requires deep understanding of the original and thorough absorption of the style and meaning, and it is this which is often lacking.

(D. Hill, personal communication, 2010)

Due to the high complexity involved in simplification, a perfect simplified cloning of the original can hardly be possible. Whatever effort to preserve the original, some distortions cannot be avoided. This is not to say simplification is not worth trying. Some simplifications in the major international series are very successful, which shows that simplification can be nicely done and simplification presents very good stories. Our first point to make here is to show the gap between simplification and the original text, and the second point to make is to frankly state the complexity involved in simplification. A perfect simplification is to some extent mission impossible, but it is always worth the effort.

2 Strategies to deal with culture-related issues

One of the benefits associated with graded readers is that it cultivates culture awareness and facilitates culture study, as the stories in graded readers are set in different cultural and social backgrounds. This also brings a problem that the readers' lack of background information of the stories may hinder their deeper understanding of the stories.

Besides adding new titles on culture and history to facilitate culture study, a

possible way to facilitate culture study is to create the internal consistence among the currently available titles in the same series, without taking the trouble to publish new titles. Currently, the common practice of displaying the titles in one series by the major international publishers is to present the titles by level of difficulty or in an alphabetical order. A few minor changes to the current practice will help to establish internal consistency within a series to facilitate culture study. The first way to establish the internal consistency within one series is to classify the books written by the same author into one category. For example, in the OBW series there are several books written by Charles Dickens: *A Christmas Carol* at Level 3, *A Tale of Two Cities* at Level 4, *Great Expectation* at Level 5 and *David Copperfield* at Level 5. Books of the same authorship may create a new module for culture study, for example in the Dickens's case, the British society in the nineteenth century. The second way to establish the internal consistency within the same series is to set the links between related titles. For example, *The Bronte Family* at OBW Level 3 is an optimal supplementary reader to *The Wuthering Height* at OBW Level 5 and *Jane Eyre* at OBW Level 6. Students may better understand *The Wuthering Height* and *Jane Eyre* through *The Bronte Family*, e.g., the circumstance within which *The Wuthering Height* and *Jane Eyre* was written, background knowledge of governess and boarding school in the British society, etc.

Another suggestion to deal with the culture-related issues is to remove the titles which may not be appropriate for a particular context. Take the present study as an example. The participants in the present study were aged from 15 to 17. In mainland China, children and teenagers under 18 are described as flowers of the country even in government documents. They are supposed to be pure and delicate. The parents and education authority do not expect them to be exposed to some sensitive topics such as violence and sex. One grade reader, which is adapted from Edgar Ellen Poe, tells a story that a mad owner cut one of its cat's eyes from its socket alive. This content is likely to subject extensive reading to challenge and criticism, because the reading material contains something negative. It is always one thing that the students actually know violence and sex, but it is another thing to include these in an extensive reading program. It is the coordinator's responsibility to ensure that an extensive reading program is

successful and fruitful, and at the same time to make it safe, far from any unnecessary criticism or question.

3 Efforts to release the lexical burden

The first important factor to ensure a smooth reading experience is to release the lexical burden of the students in their reading process. According to the researcher's observation, in some cases the lexical burden or the use of unknown words cannot be avoided because of the development of the plot; rather, in some cases the use of unknown words can be avoided. The first situation in which the use of unknown words can be avoided has been discussed in Section 4.4.4, that is the introduction of the non-standard forms.

In addition to the unknown words in the body text, the second source of unknown words comes from the glossary and footnote, in which some difficult words are explained in English. However, in some cases the words used to explain a difficult word are even more difficult than the original one. For example, in *The Christmas Carol* published by Black Cat, some Christmas ghosts are introduced. It is a good piece of supplementary reading to this story except for some very difficult words. Some difficult words are footnoted to aid comprehension. In one footnote, *skull* is explained as *craniums*. This footnote does not work since students know neither of them. Using more difficult words to explain an unknown word happens often. It deserves the editors' attention that difficult words in glossaries and footnotes will also cause difficulty to students in extensive reading.

The third type of lexical burden is caused by the publishers' unfamiliarity with a particular context. For example, in a *Christmas Carol*, there are two sentences describing Scrooge's character:

He was a miser. And he was a cold and solitary man.

Interestingly, in this book *miser* and *solitary* are not footnoted but *business partners* and *mean man* are. For students receiving English education in mainland China, *miser* and *solitary* are more difficult to understand. If the books

are intended for the EFL learners in mainland China, it is *miser* and *solitary* that deserve their footnotes given the messages they carry. The explanation for this may be that the major international graded reader series are intended for the learners of English all over the world instead of a particular country. In different countries, the words causing difficulty to readers are different, while the publishers footnote or gloss the words which are considered to be generally difficult. This generalization sometimes causes problems to readers in a particular country, and some of them considered that the editors did not have an accurate knowledge of their vocabulary. To avoid the readers' disappointment, the publishers are advised to make slight modification in terms of footnotes and glossary to ensure that the words chosen for further explanation are those known to the students. A comparison between the graded readers and the local English language teaching syllabus will work.

References

- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Alderson, J. C., & Urquhart, A. H. (1984). What is reading? In J. C. Alderson & A. H. Urquhart (Eds), *Reading in a foreign language* (pp. xv-xxviii). London and New York: Longman.
- Alessi, S., & Dwyer, A. (2008). Vocabulary assistance before and during reading. *Reading in a Foreign Language*, 20(2), 246-263.
- Al-Homoud, F., & Schmitt, N. (2009). Extensive reading and a challenging environment: A comparison of extensive reading and intensive reading approaches in Saudi Arabia. *Language Teaching Research*, 13, 383-401.
- Allan, R. (2009). Can a graded reader corpus provide “authentic” input? *ELT Journal*, 63(1), 23-32.
- Al-Rajhi, A. (2004). *Joining the online literacy club: Internet reading among Saudi EFL learners*. Unpublished Ph.D. dissertation, Indiana University of Pennsylvania, USA, Pennsylvania.
- Alshamrani, H. M. (2003). *The attitudes and beliefs of ESL students about extensive reading of authentic texts*. Unpublished Ph.D dissertation, Indiana University of Pennsylvania, Pennsylvania, USA.
- Alshwairkh, S. A. N. (2004) *Learning vocabulary through Internet reading: Approaches and attitudes of ESL MBA students*. Unpublished Ph.D. dissertation, Indiana University of Pennsylvania, Pennsylvania, USA.
- Alyousef, H. S. (2005). Teaching reading comprehension to ESL/EFL learners. *The Reading Matrix*, 5 (2), 143-154.
- Ambatchew, M. D. (2003). *The effect of primary English readers on reading skills in Ethiopia: A study in African educational needs*. Unpublished Ph.D dissertation, University of Pretoria, Pretoria.
- An, M. (2010). Use of communication strategies by Chinese EFL learners. *Chinese Journal of Applied Linguistics*, 33 (3), 110-125.
- Anderson, J. R. (1995). *Cognitive psychology and its implications* (2nd edition). New York: Freeman.
- Ariyanto, S. (2009). Encouraging extensive reading to improve academic literacy

- in the EFL class. In A. Cirocki (EdS.), *Extensive reading in English language teaching* (pp. 295-309). Munich, Germany: Lincom.
- Arnold, N. (2009). Online extensive reading for advanced foreign language learners: An evaluation study. *Foreign Language Annals*, 42(2), 340-366.
- Asraf, R. M., & Ahmad, I. S. (2003). Promoting English language development and the reading habit among students in rural schools through the Guided Extensive Reading program. *Reading in a Foreign Language*, 15 (2), 83-102.
- Author, T., & Stevens, V. (1996). A principled consideration of computers and reading in a second language. In M. Pennington (Eds.), *The power of CALL* (pp. 115-136). Houston: Athelstan.
- Bacon, S., & Finnemann, M. (1992). Sex differences in self-reported beliefs about foreign-language learning and authentic oral and written input. *Language Learning*, 42, 471-495.
- Baddeley, A. (2006). Working memory: An overview. In S. Pickering (eds.), *Working memory and education* (pp. 1-31). Burlington, MA: Academic Press.
- Baddeley, A. (2007). *Working memory, thought and action*. New York: Oxford University Press.
- Baker, C. (2006). *Foundations of bilingual education and bilingualism* (4th edition). Clevedon, UK: Multilingual Matters.
- Bamford, J. (1984). Extensive reading by means of graded readers. *Reading in a Foreign Language*, 2, 218-260.
- Bamford, J., & Day, R. R. (1997). Extensive reading: What is it? Why bother? *The Language Teacher*, 21(5), 6-8.
- Bamford, J., & Day, R. R. (2004). *Extensive reading activities for teaching language*. Cambridge: Cambridge University Press.
- Baker, C. L., & McCarthy, J. J. (1981). *The logical problems of language acquisition*. Cambridge, Mass: MIT Press.
- Barrett, M. E., & Datesman, M. K. (1992). *Reading on your own: An extensive reading course*. Boston: Heinle & Heinle.
- Barton, D. (1994). *Sustaining local literacies*. Clevedon: Multilingual Matters.
- Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography*, 6, 253-279.

- Bell, T. (2001). Extensive reading: Speed and comprehension. *The Reading Matrix, 1* (1), 1-13.
- Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change? *The International Journal of Bilingual Education and Bilingualism, 10*, 210-223.
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12*: 3-11.
- Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading, 9*, 43-61.
- Bialystok, E., Martin, M. M., & Viswanathan, M. (2005). Bilingualism across the lifespan: The rise and fall of inhibitory control. *International Journal of Bilingualism, 9*, 103-119.
- Bialystok, E., McBride, C. C., & Luk, G. (2005). Bilingualism, language proficiency, and learning to read in two writing systems. *Journal of Educational Psychology, 97*, 580-590.
- Blau, E. K. (1982). The effect of syntax on readability for ESL students in Puerto Rico. *TESOL Quarterly, 16*(4), 517-528.
- Bright, J. A., & McGregor, G. P. (1970). *Teaching English as a second language*. London: Longman.
- Brooks, N. (1960). *Language and language learning*. New York: Harcourt, Brace & World.
- Broughton, G., Brumfit, C., Flavel, R., Hill, P., & Pincas, A. (1978). *Teaching English as a foreign language*. London: Routledge & Kegan Paul.
- Brown, D. (1994). *Principles of language learning and teaching*. NJ: Prentice Hall.
- Brown, D. (2008). Why and how textbooks should encourage extensive reading. *ELT Journal, 63*(3), 238-245.
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening. *Reading in a Foreign Language, 20*(2), 136-163.
- Bruton, A. (2002). Extensive reading is reading extensively, surely? *The Language Teacher, 26* (11), 23-25.
- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.

- Byrne, D. (1986). *Teaching oral English*. Harlow: Longman.
- Cain, K. (2006). Children's reading comprehension: The role of working memory in normal and impaired development. In S. Pickering (eds.), *Working memory in education* (pp. 61-91). Burlington, MA: Academic Press.
- Cambourne, B. (2002). Holistic, integrated approaches to reading and language arts instruction: The constructivist framework of an instructional theory. In A. E. Farstrup, & S. Samuels (Eds.), *What research has to say about reading instruction* (pp. 25-47). Newark, DE: International Reading Association.
- Campbell, T., & Stanley, C. (1966). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally & Co.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.
- Careless, D. R. (2009). Revisiting the TBLT versus P-P-P debate: Voices from Hong Kong. *Asian Journal of English Language Teaching*, 19, 49-66.
- Carlisle, J. F., Beeman, M., Davis, L. H., & Spharim, G. (1999). Relationship of meta-linguistic capabilities and reading achievement for children who are becoming bilingual. *Applied Psycholinguistics*, 20, 459-478.
- Carrell, P. L. (1992). Awareness of text structure: Effects on recall. *Language Learning*, 42, 1-20.
- Carroll, J. B., Davies, P., & Richman, B. (1971). *The American heritage word frequency book*. New York: American Heritage Publishing Co.
- Caruso, J. M. (1994). *The effects of extensive reading on reading comprehension and writing proficiency in foreign language learning*. Unpublished MA dissertation. West Virginia University, West Virginia, USA.
- Carver, R. P. (1992). Effect of prediction activities, prior knowledge, and text type on the amount of comprehension: Using rauding theory to critique schema theory research. *Reading Research Quarterly*, 27, 164-174.
- Carver, R. P. (1994). Percentage of unknown words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior*, 26, 413-437.
- Carver, R. P. (2003). The highly lawful relationships among pseudoword decoding, word identification, spelling, and reading. *Scientific Studies of*

- Reading*, 7, 127-154.
- Celce-Murcia, M. (2001). Language teaching approaches: An overview. In M. Celce-Murcia (Ed.) *Teaching English as a second or foreign language* (pp. 3-11). Boston: Heinle & Heinle.
- Celce-Murica, M, & Hilles, S. (1988). *Techniques and resources in teaching grammar*. Oxford: Oxford University Press.
- Cheng, H. F. (2004). A comparison of multiple-choice and open-ended response formats for the assessment of listening proficiency in English. *Foreign Language Annals*, 37 (4), 544-553.
- Cheng, L. Y. (2008). The key to success: English language testing in China. *Language Testing*, 25 (1), 15-37.
- Cheng, L. Y., & Qi, L. X. (1996). Test review: description and examination of the National Matriculation English Test. *Language Assessment Quarterly*, 3 (1), 53-70.
- Chew, L. (2003). *Extensive reading in English: Hong Kong secondary one students' response to high-interest unsimplified stories*. Unpublished MA dissertation, The University of Hong Kong, Hong Kong.
- Cho, K. S., & Krashen, S. D. (1994). Acquisition of vocabulary from the Sweet Valley Kids series: Adult ESL acquisition. *Journal of Reading*, 37, 662-667.
- Cho, K. S., & Krashen, S. D. (2001). Sustained silent reading experiences among Korean teachers of English as a foreign language: The effect of a single exposure to interesting, comprehensible reading. *Reading Improvement*, 38(4), 170-175.
- Clark, C., & Rumbold, K. (2006). Reading for pleasure: A research overview. National Literacy Trust. Available at: <http://www.literacytrust.org.uk/search?q=Reading+for+pleasure%3A+A+research+overview>.
- Claridge, G. (2005). Simplification in graded readers: Measuring the authenticity of graded texts. *Reading in a Foreign Language*, 17 (2), 144-158.
- Clay, M. (1993). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- Coady, J. (1997). L2 vocabulary acquisition through extensive reading. In Coady, J., & Huckin, T. (Eds). *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 225-237). Cambridge: Cambridge University

Press.

- Coady, J., & Huckin, T. (1997). (eds.) *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge: Cambridge University Press.
- Cobb, T. (2005). The case for computer-assisted extensive reading. *Contact, Special Research Symposium*, 31(2), TESOL Ontario, 2005. available at: http://www.er.uqam.ca/nobel/r21270/cv/comp_assist_er.htm.
- Cohen, A. D. (1995). In which language do/should multilingual think? *Language, Culture and Curriculum*, 8(2), 99-113.
- Cohen, A. D. & Manion, L. (1994). *Research methods in education* (4th edition). London/New York: Routledge.
- Coleman, A. (1930). A new approach to practice in reading a modern language. *The Modern Language Journal*, 15(2), 101-118.
- Coley, J. D. (1983). *Project READ: Observations from the past and implications for the future*. ERCI document No. 243363, available at: http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=ED234363&_ERICExtSearch_SearchType_0=no&accno=ED234363.
- Connolly, P. (2007). *Quantitative data analysis in education: A critical introduction using SPSS*. New York: Routledge.
- Constantino, R. (1995). Learning to read in a second language doesn't have to hurt: The effect of pleasure reading. *Journal of Adolescent and Adult Literacy*, 39, 68-69.
- Cooper, M. (1984). Linguistic competence of practiced and unpracticed non-native readers of English. In J. C. Alderson & A. H. Urquhart (Eds), *Reading in a foreign language* (pp. 122-138). London and New York: Longman.
- Coxhead, A. (1998). *An Academic Word List*. Occasional Publication Number 18, LALS, Victoria University of Wellington, New Zealand.
- Cummins, J. (1991). Conversational and academic language proficiency in bilingual contexts. *AILA Review*, 8, 75-89.
- Davies, A. (1984). Simple, simplified and simplification: What is authentic? In J. C. Alderson, & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 181-195). London: Longman.
- Davis, C. (1995). Extensive reading: An expensive extravagance? *ELT Journal*

48(4): 329-336.

- Davis, F. B. (1968). Research in comprehension in reading. *Reading Research Quarterly*, 3, 499-545.
- Davis, J. N., & Bistodeau, L. (1993). How do L1 and L2 reading differ? Evidence from think aloud protocols. *Modern Language Journal*, 77, 459-472.
- Davidson, P., & Williams, E. (2005). *Outcome evaluation of extensive reading programmes in Africa*. IATEFL 2005 Cardiff Conference Selections, 76-77.
- Day, R. R. (2004). Two writing activities for extensive reading. *English Teaching Forum*, 42(3), Available at: <http://eca.state.gov/forum/vols/voLevel42/no3/p8.htm>.
- Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.
- Day, R. R., & Bamford, J. (2000). Reaching reluctant readers. *English Teaching Forum*, 38(3), 12-17.
- Day, R. R., & Bamford, J. (2002). Top ten principles for teaching extensive reading. *Reading in a Foreign Language*, 14, 375-411.
- Day, R. R., & Hirsch, K. (2004). Tips for reading extensively. *English Teaching Forum*, 42(4), Available at: <http://eca.state.gov/forum/vols/voLevel42/no4/p12.htm>.
- Day, R. R., Omura, C., & Hiramatsu, M. (1991). Incidental EFL vocabulary learning and reading. *Reading in a Foreign Language*, 7(2), 541-551.
- De' Ath, P. (2001). The Niue literacy experiment. *International Journal of Educational Research*, 35 (2), 137-146.
- Deckert, G. (2006). What helped highly proficient EFL learners the most? *TESL Reporter*, 39(2), 1-15.
- de Morgado, N. F. (2009). Extensive reading: Students' performance and perception. *The Reading Matrix*, 9 (1), 31-43.
- Dupuy, B., & Krashen, S. D. (1993). Incidental vocabulary acquisition in French as a foreign language. *Applied Language Learning*, 4 (1), 55-64.
- Dwyer, E., & West, R. (1994). Effects of silent sustained reading on reading rate among college students. (ERIC Documentation Reproduction Service No. ED382924). Available at: <http://eric.ed.gov/PDFS/ED382924.pdf>.

- Elley, W. B. (1987). Evaluating the benefits of Book Floods in second language learning. In Bickely, V.B. (eds), *Further Directions in English Language Teacher Education: Asia and Pacific perspectives*, (pp. 221-225). Hong Kong: Institute of Language in Education.
- Elley, W. B. (1989). Vocabulary acquisition form listening to stories. *Reading Research Quarterly*, 24 (2), 174-187.
- Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning*, 41, 375-411.
- Elley, W. B. (2000). The potential of book floods for raising literacy levels. *International Journal of Education*, 46, 233-55.
- Elley, W. B., & Foster, D. (1996). *Sri Lanka Books in Schools Project: Final report*. London: International Book Development.
- Elley, W. B., & Mangubhai, F. (1981). *The impact of a book flood in Fiji primary schools*. Wellington: New Zealand Council for Educational Research.
- Elley, W. B., & Mangubhai, F. (1983). The impact of reading on second language learning. *Reading Research Quarterly*, 19, 53-76.
- Ellis, N. C. (1994). Implicit and explicit language learning: An overview. In N. C. Ellis (Eds.), *Implicit and explicit learning of languages* (pp.1-31). London: Academic Press.
- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305-352.
- Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Basil Blackwell.
- Ellis, R. (1994). *The study of second language acquisition*. New York: Oxford University Press.
- Enright, M., Grabe, W., Koda, K., Mosenthal, P., & Mulcahy-Ernt, P. (2000). TOEFL 2000 reading framework: A working paper. Princeton, NL: Educational Testing Service.
- Eskey, D. E. (2002). Reading and the teaching of L2 reading. *TESOL Journal*, 11(1), 5-9.
- Ewert, D. E. (2009). Making connections: Using literature for extensive reading. In A. Cirocki (Ed.), *Extensive reading in English language teaching* (pp. 387-407). Munich, Germany: Lincom.
- Flahive, D. E., & Bailey, N. H. (1993). Exploring reading/writing relationships in

- adult second language learners. In J. E. Carson, & I. Leki, (eds.), *Reading in the composition classroom: Second language perspectives* (pp. 128-140). Boston, MA: Heinle and Heinle.
- Flesh, R., & Lass, A. H. (1996). *The classic guide to better writing*. New York: Harper Perennial.
- Florit, E., Roch, M., Altoe, G., & Levorato, M.C. (2009). Listening comprehension in preschoolers: The role of memory. *British Journal of Developmental Psychology*, 27, 935-951.
- Fredricks, L., & Sobko, V. (2008). Culturally relevant extensive reading in Tajikistan. *Central Eurasian Studies Review*, 7(1), 34-39.
- Fries, C. C. (1945). *Teaching and learning English as a foreign language*. Ann Arbor: University of Michigan Press.
- Fry, E. B. (1963). *Teaching faster reading: A manual*. Cambridge: Cambridge University Press.
- Fry, E. B. (1975). *Reading drills for speed and comprehension* (2nd ed.). Providence, RI: Jamestown Publishers.
- Fu, K. (1984). *Zhongguo waiyu jiaoxueshi [History of foreign language teaching in China]*. Shanghai: Shanghai Foreign Language Education Press.
- Fujigaki, E. (2009). Extensive reading for weak readers: Developing reading fluency in the EFL/ESL context. In A. Cirocki (Eds.), *Extensive reading in English language teaching* (pp. 273-293). Munich, Germany: Lincom.
- Fujimoto-Adamson, N. (2006). Globalization and history of English education in Japan. *The Asian EFL Journal Quarterly*, 8 (3), 259-282.
- Furukawa, A. (2005). *The complete book guide for extensive reading*. Tokyo: Cosmopier Publishing.
- Furukawa, A. (2008). Extensive reading from the first day of English learning. *Extensive reading in Japan*, 1(2), 11-14.
- Gao, S. (2008). *An empirical study of applying 21st Century to college English extensive reading teaching*. Unpublished MEd thesis, Central China Normal University, Wuhan, China.
- Gardener, D. (2008). Vocabulary recycling in children's authentic reading material: A corpus-based investigation of narrow reading. *Reading in a foreign language*, 20 (1), 92-122.
- Garrett, N. (1986). The problem with grammar: What kind can the language

- learner use? *The Modern Language Journal*, 70, 133-148.
- Gass, M. S., & Selinker, L. (2008). *Second language acquisition: An introductory course (3rd edition)*. New York: Routledge.
- Gee, R. W. (1999). Encouraging ESL students to read. *TESOL Journal*, 8(1), 3-7.
- Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes*, 25, 265-304.
- Goodman, K. S. (1982). *Language and literacy: The selected writings of Kenneth S. Goodman (1): Process, theory, research*. Boston: Routledge & Kegan Paul.
- Gorsuch, G., & Taguchi, E. (2009). Repeated reading and its role in an extensive reading programme. In A. Cirocki (ed.), *Extensive reading in English language teaching* (pp. 249-272). Munich, Germany: Lincom.
- Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, 25 (3), 375-406.
- Grabe, W. (2004). Research on teaching reading. *Annual Review of Applied Linguistics*, 24, 44-69.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.
- Grabe, W., & Stoller, F. L. (1996). Reading and vocabulary development in a second language: A case study. In J. Coady, & T. Huckin (Eds). *Second language vocabulary acquisition: A rationale for pedagogy*. Cambridge: Cambridge University Press.
- Grabe, W., & Stoller, F. L. (2002). *Teaching and researching reading*. Harlow: Longman.
- Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading (2nd edition)*. Harlow: Longman.
- Graves, M. F. (2006). *The vocabulary book: Learning and instruction*. New York: Teachers College Press.
- Graves, M. F., & Watts-Taffe, S. (2002). The place of word consciousness in a research-based vocabulary program. In A. Farstrup & S. Samuels (eds.), *What research has to say about reading instruction (3rd edition)* (pp. 140-165). Newark, DE: International Reading Association.
- Green, C. (2005). Integrating extensive reading in the task-based curriculum. *ELT Journal*, 59 (4), 306-311.

- Greenberg, D., Rodrigo, V., Berry, A., Brinck, T., & Joseph, H. (2006). Implementation of an extensive reading program with adult learners. *Adult Basic Education*, 16(2), 81-97.
- Greenwood, J. (1988). *Class readers*. Oxford: Oxford University Press.
- Griffiths, C., & Parr, J. M. (2001). Language-learning strategies: Theory and perception. *ELT Journal*, 53 (3), 247-254.
- Gu, H. P. (2009). Qianxi jiaoji jiaoxuefa zai zhongguo de fazhan xianzhuang [A preliminary analysis of the development and current situation of communicative language teaching in China]. *Wenjiao Ziliao*, 19, 56-57.
- Gu, Y. Q., & Dong, L. Z. (2005). Xianggang shuangyu jiaoxue changshi de jingyan ji qishi' [Lessons from Kong Kong's experiment with English-medium instruction]. *Xiandai Waiyu [Modern Foreign Languages]*, 28(1), 43-52.
- Guan, Q., & Meng, W. J. (2007). China's new national curriculum reform: Innovation, challenges and strategies. *Frontiers of Education in China*, 2(4), 579-604.
- Guthrie, J. T. & Wigfield, A. (2000). Engagement and motivation in reading. In M. Kamil, P. Mosenthal, P. D. Pearson, and R. Barr (eds), *Handbook of reading research*, Vol III, (pp. 403-422). Mahwah, NJ: Lawrence Erlbaum.
- Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., Scaffidi, N. T., & Tonks, S. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology*, 96(3), 403-423.
- Guo, N. H. (2007). *Graded readers and young adult literature in an extensive reading classroom*. Unpublished MA thesis, National Tsinghua University, Taiwan.
- Hafiz, F., & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal*, 43, 4-13.
- Hair, J. F., Anderson, R. E., Tathan, R. L., & Black, W. C. (1998). *Multivariate data analysis* (5th edition). New Jersey: Prentice-Hall.
- Hamp-Lyons, E. (1983). Developing a course to teach extensive reading skills to university-bound ESL learners. *System*, 11, 303-312.
- Han, E. J. (2010). 2010 gaokao shijuan fenxi [Analysis of the 2010 National Matriculation English Test]. A talk delivered to English teachers in

- Jinshan District, Shanghai, in Teacher Training College, Jinshan District, Shanghai.
- Han, Z. H., Anderson, N. J., & Freeman, D. (2009). Introduction: Crossing the boundaries. In Z. H. Han & N. J. Anderson (eds.), *Second language reading research and instruction: Crossing the boundaries*. Ann Arbor: The University of Michigan Press.
- Hang, B. T. (1993). *Zhongxue yingyu jiaoxuefa [English language teaching methods for secondary school]*. Shanghai: East China Normal University Press.
- Hang, B. T. (2000). *Zhongxue yingyu jiaoxuefa [English language teaching methods for secondary school] (2nd edition)*. Shanghai: East China Normal University Press.
- Harmer, J. (2007). *The practice of English language teaching* (4th edition). Harlow: Longman.
- Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for teaching teachers. *The Reading Teacher*, 59, 636-644.
- Hayashi, K. (1999). Reading strategies and extensive reading in EFL classes. *RELC Journal*, 30 (2), 114-132.
- Hazenburgh, S., & Hulstijn, J. H. (1996). Defining a minimal receptive second language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, 17, 145-163.
- Heaton, J. B. (1975). *Writing English language tests: a practical guide for teachers of English as a second or foreign language*. London: Longman.
- Hedge, T. (1985). *Using readers in language teaching*. London: Macmillan.
- Heidi. (2010). Optimal silent & oral reading rates for reading fluency: Graded reading levels. A Reading Horizon Post. Available at: <http://www.readinghorizons.com/blog/author/Heidi.aspx>.
- Helgesen, M. (1997). Bringing those books back to the classroom: Tasks for extensive reading. *The Language Teacher*, 21(5), 53-54.
- Helgesen, M. (2005). Extensive reading reports-different intelligences, different levels of processing. *Asian EFL Journal*, 7(3), Article 2.
- Hickey, T. (1991). Leisure reading in a second language: An experiment with audio-tapes in Irish. *Language, Culture and Curriculum*, 4(2), 119-131.
- Higgins, J., & Wallace, R. (1989). Hopalong: A computer reader pacer. *System*,

- 17 (3), 389–399.
- Hill, D. R. (1983). *Description of a programme of extensive reading*. Handout, 2 pages. University of Edinburgh, UK.
- Hill, D. R. (1997a). Graded readers. *ELT Journal*, 51, 57-79.
- Hill, D. R. (1997b). Setting up an extensive reading programme: Practical tips. *The Language Teacher*, 21 (5), 17-20.
- Hill, D. R. (2008). Graded readers in English. *ELT Journal*, 62(2), 184-204.
- Hill, D. R., & Thomas, H. R. (1988a). Survey review: Graded readers I. *ELT Journal* 42(1): 44-52.
- Hill, D. R., & Thomas, H. R. (1988b). Survey review: Graded readers II. *ELT Journal* 42(2), 124-136.
- Hirsh, D., & Nation, I.S.P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8, 689-696.
- Holdaway, D. (1979). *The foundations of literacy*. Sydney: Ashton Scholastic.
- Holden, J. (2004). *Creative Reading*. London: Demos.
- Honeyfield, J. (1977). Simplification. *TESOL Quarterly*, 11, 431-440.
- Horst, M. (2000). *Text encounters of the frequent kind: Learning L2 vocabulary through reading*. Unpublished Ph.D thesis. University of Wales, Swansea, UK.
- Horst, M. (2005). Learning L2 vocabulary through extensive reading: a measurement study. *The Canadian Modern Language Review*, 61 (3), 355-382.
- Horst, M. (2009). Developing definitional vocabulary knowledge and lexical access speed through extensive reading. In Z. H. Han & N. J. Anderson (eds), *Second language reading research and instruction: Crossing the boundaries*, (pp. 40-64). Ann Arbor, MI: University of Michigan Press.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond a Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11, 207-223.
- Horst, M., & Meara, P. (1999). Test of a model of predicting second language lexical growth through reading. *The Canadian Modern Language Review*, 56, 308-328.
- Hsui, V. Y. (1994). A modified sustained silent reading programme for secondary

- classrooms. In S. E. A. Lim, M. Siripathy, & V. Saravan (Eds.), *Literacy: Understanding the learners needs* (pp. 165-174). Singapore: Singapore Society for Reading and Literacy.
- Hu, G. W. (2002). Recent important developments in secondary English-language teaching in the People's Republic of China. *Language, Culture and Curriculum*, 15, 30-49.
- Hu, G. W. (2005a). English language education in China: Policies, progress, and problems. *Language Policy*, 4, 5-24.
- Hu, G. W. (2005b). Reforms of basic English-language education in China: An overview. *International Journal of Educational Reform*, 14, 140-65.
- Hu, G. W. (2008). The misleading academic discourse on Chinese-English bilingual education in China. *Review of Educational Research*, 78, 195-231.
- Hu, G. W. (2009). The craze for English-medium education in China: Driving forces and looming consequences. *English Today*, 25 (4), 47-54.
- Hu, G. W., & Alsagoff, L. (2010). A public policy perspective on English medium instruction in China. *Journal of Multilingual and Multicultural Development*, 31(4), 365-382.
- Hu, M. T., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430
- Hu, W. Z. (2006). *Huwenzhong yingyu jiaoyu zixuanji*. Beijing: Foreign Language Teaching and Research Press.
- Huang, J. B., Chen, Y. H., Xu, Y., Li, J., & Fu, Y. (2004). Daxue yingyu kecheng jiaoxue yaoqiu cibiao xiuding tantao [On the vocabulary size of the College English Syllabus]. *Wai Yu Jie [Foreign Language World]*, 99, 2-9.
- Huang, R. J., & Feng, Y. Z. (1984). Gaoxiao wenli gongke xinsheng yingyu shuiping diaocha fenxi [An investigation and analysis of the English language proficiency of the freshmen in art, science and engineering]. *Foreign Language*, 5, 60-64.
- Huckin, T., & Coday, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 21, 181-193.
- Hudson, T. (2007). *Teaching second language reading*. Oxford: Oxford University Press.
- Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading

- vocabulary. *Reading in a Foreign Language*, 17 (1), 23-59.
- Hyland, K. (1990). Purpose and Strategy: Teaching extensive reading skills. *English Teaching Forum*, 28 (2), 14-18.
- Hyland, K. (1997). Is EAP necessary? A survey of Hong Kong undergraduates. *Asian Journal of English Language Teaching*, 7, 77-99.
- In'nami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing*, 26 (2), 219-244.
- Institute of Curriculum and Textbook. (1999). *Ershi shiji zhongguo zhongxiaoxue kecheng biao zhun. jiaoxue dagang huibian: waiguoyu juan (yingyu)* [Curriculum standards for Chinese primary and secondary school in the 20th century: Compilation of syllabuses: Foreign language volume]. Beijing: People's Education Press.
- Irvine, A. (2006). *Extensive reading and L2 development: A study of Hong Kong secondary learners of English*. Unpublished Ph.D dissertation. University of Edinburgh, UK.
- Iwahori, Y. (2008), Developing reading fluency: A study of extensive reading in EFL. *Reading in a Foreign Language*, 20(1), 70-91.
- Jackson, K. (2005). Rewarding reading. *English Teaching Professional*, 39, 15-18.
- Jacobs, G. M. (1991). *Second language reading recall as a function of vocabulary glossing and selected other variables*. Unpublished Ph.D dissertation, University of Hawai'i, Manoa.
- Jacobs, G. W., Renandya, W. A., Robb, T. N., Chau, M. H., Waring, R., Loh, J. K. K., Viney, P., Hassan, F., & Miles, S. (2007 to now). *Annotated bibliography of works on extensive reading in a second language*. Available at: <http://erfoundation.org/bib/biblioall.php?sortby=date>.
- Janopoulos, M. (1986). The relationship of pleasure reading and second language writing proficiency. *TESOL Quarterly*, 20 (4), 763-768.
- Janopoulos, M. (2009). Pleasure reading and writing in a second language: How can we make the connection? In A. Cirocki (Ed.), *Extensive reading in English language teaching* (pp. 429-438). Munich, Germany: Lincom.
- Johnson, R. K. & Yau, A. S. N. (1996). Coping with second language texts: The development of lexically-based reading strategies. In D.A. Watkins, & J. B.

- Biggs (eds.), *The Chinese Learner: Cultural psychological and contextual influences*. Hong Kong: Comparative Education Research Centre, Faculty of Education, University of Hong Kong.
- Jensen, L. (1986). Advanced reading skills in a comprehensive course. In F. Dubin, D. E. Eskey, & W. Grabe (Eds.), *Teaching second language reading for academic purposes* (pp. 103-124). Reading, MA: Addison-Wesley.
- Jia, G. J. (1996). *Psychology of foreign languages education*. Guangxi: Guangxi Education Press.
- Jin, K., & Zhuang, Y. X. (2002). Shuangyu jiaoxue zheng liaoyuan [Bilingual education is spreading like a wild fire]. *Jiefang Daily*, March 4.
- Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Karp, A. S. (2002). *Modification of glosses and its effect on incidental L2 vocabulary learning in Spanish*. Unpublished doctoral dissertation, University of California, Davis, USA.
- Kelly, L. G. (1969). *25 centuries of language teaching: An inquiry into the science, art, and development of language teaching methodology*, 500 B.C.-1969. Rowley, MA: Newbury House Publishers.
- Kennedy, M. L. (1985). The composing processes of college students writing from sources. *Written Communication*, 2, 434-56.
- King, J. K. (1969). A reading program for realists. *The German Quarterly*, 42(1), 65-80.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-complementary model. *Psychological Review*, 95, 163-182.
- Kitao, K., Yamamoto, M., Kitao, S. K., & Shimatani, H. (1990). Independent reading in English: Use of graded readers in the library English as a second language corner. *Reading in a Foreign Language*, 6, 383-398.
- Klapper, J. (1992). Preliminary considerations for the teaching of FL reading. *Language Learning*, 6, 53-56.
- Koda, K. (2005). *Insight in second language reading: A cross-linguistic approach*. Cambridge: Cambridge University Press.
- Koda, K. (2007). Reading and language learning: Crosslinguistic constraints on second language reading development. *Language Learning Supplement*, 57, 1-44.

- Koda, K., & Reddy, P. (2008). Cross-linguistic transfer in second language reading. *Language Teaching*, 41, 497-508.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Oxford University Press.
- Krashen, S. D. (1982). Principles and practice in second language acquisition. Oxford: Pergamon Press.
- Krashen, S. D. (1988). Do we learn to read by reading? The relationship between free reading and reading ability. In D. Tannen (Eds.), *Linguistics in context* (pp. 269-298). Norwood, NJ: Ablex.
- Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73 (4), 440-464.
- Krashen, S. D. (1993). *The power of reading: insights from the research*. Englewood: Libraries Unlimited CO.
- Krashen, S. D. (2004a). *The power of reading: insights from the research*. Westport: Conn Libraries Unlimited; Portsmouth, NH: Heinemann.
- Krashen, S. D. (2004b). Applying the Comprehension Hypothesis: Some suggestions. Paper presented at 13th International Symposium and Book Fair on Language Teaching (English Teachers Association of the Republic of China), Taipei, Taiwan. Available at: http://sdkrashen.com/articles/eta_paper/all.html.
- Krashen, S. D. (2007). Free voluntary web-surfing. *The International Journal of Foreign Language Teaching*, 3(1), 2-9.
- Krashen, S. D., & Terrell, T. (1983). *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon.
- Kucera, H., & Francis, W. N. (1967). *A computational analysis of present-day American English*. Boston: Houghton Mifflin.
- Kuhn, M., Schwanenflugel, P., Morris, R., Morrow, L., Woo, D., Meisinger, E., Sevcik, R., Bradley, B., & Stahl, S. (2006). Teaching children to become fluent and automatic readers. *Journal of Literacy Research*, 38, 357-387.
- Kuhn, M., & Stahl, S. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95, 3-21.
- Kurt, D. (2008). *Marking the mind: A history of memory*. Cambridge: Cambridge University Press.

- Kutiper, K. (1983). Extensive reading: A means of reconciliation. *English Journal*, 72(7), 58-61.
- Kutner, M., Greenberg, E., Jin, Y., Boyle, B., Hsu, Y.C., & Dunleavy, E. (2007). Literacy in everyday life: Result from the 2003 National Assessment of Adult Literacy. U.S. Department of Education. NCES 2007-480. Available at: <http://eric.ed.gov/PDFS/ED495996.pdf>.
- Kweon, S., & Kim, H. (2008). Beyond raw frequency: Incidental vocabulary acquisition in extensive reading. *Reading in a Foreign Language*, 20(2), 191-215.
- Lai, F. K. (1993a). The effects of a summer reading course on reading and writing skills. *System*, 21, 87-100.
- Lai, F. K. (1993b). Effect of extensive reading on English learning in Hong Kong. *CUHK Education Journal*, 21(1), 23-36.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P.J.L. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan.
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). New York: Cambridge University Press.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *The Canadian Modern Language Review*, 59(4), 567-587.
- Lautamatti, L. (1978). Observations on the development of the topic in simplified discourse. In V. Kohonen and N.E. Enkvist (Eds), *Text, linguistics, cognitive learning and language teaching*. Helsinki: Akateeminen Kirjakauppa Postilokero.
- Lee, S. Y. (2005). Sustained silent reading using assigned reading: Is comprehensible input enough? *The International Journal of Foreign Language Teaching*, 1(4), 10-12.
- Lee, S. Y. (2006). A one-year study of SSR: University level EFL students in Taiwan. *The International Journal of Foreign Language Teaching*, 2(1), 6-8.
- Lee, S. Y., & Hsu, Y. Y. (2009). Determining the crucial characteristics of extensive reading programs: The impact of extensive reading on EFL

- writing. *The International Journal of Foreign Language Teaching*, 5(1), 12-20.
- Leonard, D. (2010). *A visual approach to SPSS for Windows: A guide to SPSS 17.0*. Boston, MA: Allyn & Bacon.
- Leow, R. P. (1993). To simplify or not to simplify. *Studies in Second Language Acquisition*, 15, 333-355.
- Leow, R. P. (1997). Simplification and second language acquisition. *World Englishes*, 16, 291-296.
- Leow, R. P. (2009). Simplified written input and its effects on L2 comprehension: What the research reveals. In A. Cirocki (Ed.), *Extensive reading in English language teaching* (pp. 129-141). Munich, Germany: Lincom.
- Li, J. S. (2008). *A social constructivist model of senior high school English extensive-reading teaching*. Unpublished MEd thesis, Hebei Normal University, Sijiazhuang, China.
- Li, J. X. (2007). *Research on the teaching of extensive reading in senior high schools via Internet*. Unpublished MEd thesis, Northeast China Normal University, Lanzhou, China.
- Li, J. Y. (2007). *A feasibility study of offering English extensive reading course in junior middle schools*. Unpublished MEd thesis, Guangxi Normal University, Nanning, China.
- Li, X. (1990). How powerful can a language test be? The MET in China. *Journal of Multilingual and Multicultural Development*, 11, 393-404.
- Li, X. J. (1984). In defense of the communicative approach. *ELT Journal*, 38, 2-13.
- Li, X. P. (2007). *A new attempt at English extensive reading teaching in senior high school*. Unpublished MEd thesis, East China Normal University, Shanghai, China.
- Liem, D. H. (2005). *Using extensive reading to enhance students' perceptions and their reading ability*. Unpublished MA thesis, King Mongkut's University of Technology, Thonburi, Thailand.
- Light, T. (1970). The reading-comprehension passage and a comprehensive reading programme. *ELT Journal*, 24, 120-124.
- Lightbown, P. M. (1992). Can they do it themselves? A comprehension-based ESL course for young children. In Courchene, R., St. John, J., Trerien, C., &

- Glidden, J. (Eds.), *Comprehension-based second language teaching: Current trends* (pp. 353-370). Ottawa, ON: University of Ottawa Press.
- Lightbown, P. M., Halter, R. H., White, J. L., & Horst, M. (2002). Comprehension-based learning: The limits of “Do it yourself”. *The Canadian Modern Language Review*, 58 (3), 427-463.
- Linderholm, T., & van den Broek, P. (2002). The effect of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology*, 94, 778-784.
- Liu, A. J. (2010). On pragmatic “borrowing transfer”: Evidence from Chinese EFL learner’s compliment response behavior. *Chinese Journal of Applied Linguistics*, 33 (4), 26-44.
- Liu, D. Y. (2008a). Gaige kaifang sanshinian de zhongxiaoxue yingyu jiaocai [The thirty years’ English language teaching materials in primary and secondary schools since reform and open up]. *English teachers*, 10, 3-8.
- Liu, D. Y. (Eds.). (2008b). *Jichu waiyu jiaoyu fazhan baogao (1978-2008)* [Report of the development of basic English education (1978-2008)]. Shanghai: Shanghai Foreign Language Education Press.
- Liu, J. (2011). *Research on selection of English extensive reading materials in high school English teaching*. Unpublished MEd thesis, Wuhan University of Science and Technology, Wuhan, China.
- Liu, Y. F., & Zhao Y. Q. (2010). A study of teacher talk in interactions in English classes. *Chinese Journal of Applied Linguistics*, 33 (2), 76-86.
- Littlewood, W., & Liu, N. F. (1996). *Hong Kong students and their English*. Hong Kong: Hong Kong University/Macmillan.
- Long, D. (1986). Listening: What’s really going on in the classroom. In B. Snyder (Eds.), *Second language acquisition: Preparing for tomorrow* (pp. 28-37). Lincolnwood, IL: National Textbook Co.
- Long, D., Johns, C., & Morris, P. (2006). Comprehension ability in mature reader. In M. Traxler & M. A. Gernsbacher (eds.), *Handbook of psycholinguistics* (2nd edition) (pp. 801-833). Burlington, MA: Academic Press.
- Long, M., & Richards, J. (1998). Series preface. In C. Doughty, & J. Williams (Eds.), *Focus on form in classroom second language acquisition*, (pp. xi-xii). Cambridge: Cambridge University Press.
- Lyon, R. (1997). Statement before the Committee on Education and Workforce.

- U.S. House of Representatives. Available at: <http://mirror.apa.org/ppo-OLD/lyon.html>.
- Macalister, J. (2007). Implementing extensive reading in an EAP programme. *ELT Journal*, 62 (3), 258-256.
- MacDonell, C. (2004). Making the case for pleasure reading. *Teacher Librarian*, 31(4), 30-32.
- Mason, B., & Krashen, S. D. (1997). Extensive reading in English as a foreign language. *System*, 25, 91-102.
- Mason, B. (2004). The effect of adding supplementary writing to an extensive reading program. *International Journal of Foreign Language Teaching*, 1(1), 2-16
- Martino, M., & Block, E. (1992). Let them read books. *College ESL*, 2 (2), 12-20.
- Matsumura, M. (1987). *Eigo no Riidingu [English reading]* (2nd ed.). Tokyo: Taishukan Shoten.
- Maxim, H. H. (2002). A study into the feasibility and effects of reading extended authentic discourse in the beginning German language classroom. *The Modern Language Journal*, 86(1), 20-35.
- McGinley, W. (1992). The role of reading and writing while composing from sources. *Reading Research Quarterly*, 17, 581-90.
- McNeill, A. (1994). Hong Kong learners and English words: The formal semantic gap. *Proceedings of the International Language in Education Conference 1994*, 2, 69-86.
- Mikulecky, L. J. (2009). Using Internet-based children's and young adult literature for extensive reading in EFL instruction. In A. Cirocki (Eds.), *Extensive reading in English language teaching* (pp. 333-347). Munich, Germany: Lincom.
- Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals*, 39 (2), 276-295.
- Min, H. T. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning*, 58(1), 73-115.
- Ministry of Education. (1978). *Quanrizhi shinianzhi zhongxiaoxue yingyu*

- jiaoxue dagang (shixing caoan) [English syllabus for ten-year full-time primary and secondary schools]*. Beijing: People's Education Press.
- Ministry of Education. (1980). *Quanrizhi shinianzhi zhongxiaoxue yingyu jiaoxue dagang (shixing caoan dierban) [English syllabus for ten-year full-time primary and secondary schools (2nd edition)]*. Beijing: People's Education Press.
- Ministry of Education. (1982). *Quanrizhi liunianzhi zhongdian zhongxue yingyu jiaoxue dagang (zhengqiu yijiangao) [English syllabus for six-year full-time key secondary schools]*. Beijing: People's Education Press.
- Ministry of Education. (2000a). *Jiunian yiwu jiaoyu quanrizhi chuji zhongxue yingyu jiaoxue dagang (shiyong xiudingban) [English syllabus for nine-year compulsory education full-time junior secondary schools]*. Beijing: People's Education Press.
- Ministry of Education. (2000b). *Quanrizhi gaoji zhongxue yingyu jiaoxue dagang (shixianxiudingban) [English syllabus for full-time senior secondary schools]*. Beijing: People's Education Press.
- Ministry of Education. (2001). *Quanrizhi yiwu jiaoyu putong gaoji zhongxue yingyu kecheng biao zhun (shiyangao) [The English Curriculum Standards for the full-time senior high schools of compulsory education (trial edition)]*. Beijing: People's Education Press.
- Ministry of Education. (2003). *Gaozhong yingyu kecheng biao zhun (zhengqiu yijiangao) [The English Curriculum Standards for senior high schools]*. Foreign Language Teaching in Schools, Vol 7-10.
- Mok, R. (1994). Reading and English Acquisition Programme (REAP). In M. L. Tickoo (Ed.), *Research in reading and writing: A Southeast Asian collection* (pp. 30-40). Singapore: SEAMEO Regional Language Centre.
- Mombukagakusho. (2003). The course of study for foreign languages. Ministry of Education, Culture, Sports, Science and Technology. Available at: www.mext.go.jp/english/school/030301.htm.
- Nagy, W. E. & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19, 304-330.
- Nagy, W. E. & Herman, P. (1985). Incidental vs. instructional approaches to increasing reading vocabulary. *Educational Perspectives*, 23, 16-21.
- Nagy, W. E., Herman, P., & Anderson, R. C. (1985). Learning words from

- context. *Reading Research Quarterly*, 20, 233-253.
- Nagy, W. E. (1997). On the role of context in first- and second-language learning. In N. Schmitt & M. McCarthy (Eds.). *Vocabulary: description, acquisition and pedagogy* (pp.64-83). Cambridge: Cambridge University Press.
- Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. New York and London: Routledge.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2004). A study of the most frequent families in the British National Corpus. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language* (pp. 3-14). Philadelphia: J. Benjamins.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59-82.
- Nation, I. S. P. (2009). *Teaching ESL/EFL reading and writing*. New York: Routledge.
- Nation, I. S. P., & Deweerdt, J. P. (2001). A defense of simplification. *Prospect*, 16(3), 55-67.
- Nation, I. S. P., & Wang, K. M. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12, 355-380.
- Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmidt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). New York; Cambridge University Press.
- Nelson, N. & Calfee, R.C.(1998). The reading-writing connection viewed historically, in N. Nelson, & R.C. Calfee (Eds.), *The reading-writing connection: Ninety-seventh yearbook of the national society for the student of education—Part II*, (PP. 1-52). Chicago: The University of Chicago Press.
- Newmark, L. (1971). A minimal language teaching program. In P. Pimsleur & T. Quinn (Eds.), *The psychology of second language learning* (pp. 11-18). Cambridge: Cambridge University Press.
- Ng, S. M., & Sullivan, C. (2001). The Singapore reading and English acquisition program. *International Journal of Educational Research*, 35 (2), 157-167.

- Noriguchi, S. (2006). English education leaves much to be desired. Available at: www.asahi.com/english/Herald-asahi/TKY200609150129.html.
- Nunan, D. (2002). The impact of English as global language: Policy and planning in greater China. *The Hong Kong Journal of Applied Linguistics*, 7, 1–15.
- Nunan, D., & Bailey, K. M. (2009). *Exploring second language classroom research: A comprehensive guide*. Boston, MA: Heinle, Cengage Learning.
- Nuttall, C. (1982). *Teaching reading skills in a foreign language*. London: Heinemann Educational.
- Nuttall, C. (1996). *Teaching reading skills in a foreign language* (2nd edition). Oxford: Heinemann.
- Nyikos, M. (1990). Sex related differences in adult language learning: Socialization and memory factors. *The Modern Language Journal*, 3, 273-287.
- Ozburn, M. S. (1995). A successful high school sustained silent reading program. *English in Texas*, 26 (3), 4-5.
- Pachler, N., & Field, K. (1997). *Learning to teach modern foreign languages in the secondary school: A companion to school experience*. London & New York: Routledge.
- Pargment, M. S. (1943). What constitutes a reading knowledge of a foreign language, and how it can be acquired? *The French Review*, 17(2), 74-82.
- Paribakht, T.S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: a rationale for pedagogy* (pp. 174-200). Cambridge: Cambridge University Press.
- Park, G. P. (2004). Comparison of L2 listening and reading comprehension by university students learning English in Korea. *Foreign Language Annals*, 37 (3), 448-458.
- Parry, K. (1991). Building a vocabulary through academic reading. *TESOL Quarterly*, 25, 629-653.
- Pearson, P. D., & Tierney, R. J. (1984). On becoming a thoughtful reader. In A.C. Purves, & O.Niles (eds.), *Becoming readers in a complex society*, (pp. 143-173). Chicago: National Society for the Study of Education.
- Peng, J. X. (2004). *The impact of extensive reading on EFL learners' motivation to read: A preliminary study among English learners in senior high school*.

- Unpublished MEd thesis, Shoudu Normal University, Beijing, China.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357-383.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). The acquisition of reading comprehension skill. In M. Snowling & C. Hulme (Eds.), *The science of reading* (pp. 227-247). Malden, MA: Blackwell.
- Perfetti, C. A., van Dyke, J., & Hart, L. (2001). The psycholinguistics of basic literacy. In M. McGroarty et al. (eds.), *Annual Review of Applied Linguistics*, 21, (pp. 127-149). New York: Cambridge University Press.
- Pi, W. B. (2004). *Zhongxue shuangyu jiaoxue de shijian yusikao [Bilingual education in secondary schools and reflections.]*. Unpublished MA dissertation. Shanghai: East China Normal University.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18 (1), 1-28.
- Pilgreen, J., & Krashen, S. D. (1993). Sustained silent reading with English as a second language high school students: Impact on reading comprehension, reading frequency, and reading enjoyment. *School Library Media Quarterly*, 22, 21-23.
- Pino-Silva, J. (2006). Extensive reading through the Internet: Is it worth the while? *The Reading Matrix*, 6 (1), 85-96.
- Poulshock, J. (2010). Extensive graded reading in the liberal arts and sciences. *Reading in a Foreign Language*, 22(2), 304-322.
- Powell, S. (2005). Extensive reading and its role in Japanese high schools. *The Reading Matrix*, 5(2), 28-42.
- Pressley, M., & Harris, S. (2006). Cognitive strategies instruction: From basic research to classroom instruction. In P. Alexander & P. Winnie (Eds.), *Handbook of educational psychology* (2nd edition, pp. 265-286). Mahwah, NJ: Erlbaum.
- Prichard, C. (2008). Evaluating L2 readers' vocabulary strategies and dictionary use. *Reading in a Foreign Language*, 20(2), 216-231.
- Pulido, D. (2009). Vocabulary processing and acquisition through reading: Evidence for the rich getting richer. In Z. H. Han & N. J. Anderson (eds), *Second language reading research and instruction: Crossing the boundaries* (pp. 65-82). Ann Arbor: The University of Michigan Press.

- Pulido, D., & Hambrick, D. Z. (2008). The virtuous circle: Modeling individual differences in L2 reading and vocabulary development. *Reading in a Foreign Language*, 20 (2), 164-190.
- Qi, L.X. (2005). Stakeholders' conflicting aims undermine the washback functions of a high-stake test. *Language Testing*, 22 (2), 142-173.
- Qian, D. D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review*, 56, 282-308.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52, 513-536.
- Qian, D. D., & Schedl, M. (2004). Evaluating an in-depth vocabulary knowledge measure for assessing reading performance. *Language Testing*, 21, 28-52.
- Qin, J. H. (2008). *An empirical study on incidental vocabulary acquisition through extensive reading for college students*. Unpublished MEd thesis, Jilin University, Changchun, China.
- Ramaiah, M. (1994). Reading initiatives in Malaysia. In M. L. Tickoo. (Ed.), *Research in reading and writing: A Southeast Asian collection* (pp. 79-89). Singapore: SEAMEO Regional Language Centre.
- Rao, Z. H. (2010). Chinese students' perceptions of native English-speaking teachers in EFL teaching. *Journal of Multilingual and Multicultural Development*, 31 (1), 55-68.
- Ren, Y. Q. (2010). *Developing reading fluency on tertiary level: An insight into the feasibility and validity of online extensive reading*. Unpublished MEd thesis, Nanchang University, Nanchang, China.
- Renandya, W. A. (2007). The power of extensive reading. *RELC Journal*, 38, 133-149.
- Renandya, W. A., & Jacobs, G. M. (2002). Extensive reading: Why aren't we all doing it? In J. C. Richards, & W. A. Renadya (Eds.), *Methodology in language teaching: An analogy of current practice* (pp. 295-302). New York: Cambridge University Press.
- Renandya, W. A., Rajan, B. R. S., & Jacobs, G. M. (1999). Extensive reading with adult learners of English as a second language. *RELC Journal*, 30, 39-61.

- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Richards, J. C., Platt, J., & Platt, H. (1992). *Longman dictionary of language teaching and applied linguistics* (2nd edition). Harlow, Essex: Longman.
- Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching* (2nd edition). Cambridge: Cambridge University Press.
- Richards, J. C., & Schmitt, N. (2002). *Longman dictionary of language teaching and applied linguistics* (3rd edition). London: Person Education.
- Rivers, W. M. (1981). *Teaching foreign language skills* (2nd edition.). Chicago: University of Chicago Press.
- Robb, T. N., & Susser, B. (1989). Extensive reading vs. skills building in an EFL context. *Reading in a Foreign Language*, 5, 239-251.
- Robb, T. N. (2001). Extensive reading in an EFL environment. In J. Murphy & P. Byrd (Eds.), *Understanding the courses we teach* (pp. 218-235). Ann Arbor, MI: University of Michigan Press.
- Robb, T. N. (2009). The Reader Quiz module for extensive reading. In M. Thomas (Ed.), *Selected proceedings of the thirteenth annual JALT CALL SIG conference 2008* (pp. 109-116). Tokyo: Japan Association for Language Teaching.
- Rodrigo, V., Greenberg, D., Burke, V., Hall, R., Berry, A., Brinck, T., Joseph, H., & Oby, M. (2007). Implementing an extensive reading program and library for adult literacy learners. *Reading in a Foreign Language*, 19(2), 106-119.
- Rosszell, R. (2007). *Combining extensive reading and intensive vocabulary study in a Japanese university*. Unpublished Ph.D dissertation, Temple University Japan, Tokyo.
- Rubin, J. (1994). A review of second language listening comprehension research. *Modern Language Journal*, 78, 199-221.
- Rutherford, W. S. (1987). *Second language grammars: Learning and teaching*. New York: Longman.
- Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading and spelling as predictors of school achievement and choice of secondary education. *Learning and Instruction*, 18 (2), 201-210.
- Saleem, B. A. A. (2010). *Impact of extensive reading on literacy perceptions and on EFL writing quality of English major students at the Islamic University*

- of Gaza*. Unpublished MA dissertation, The Islamic University, Madinah.
- Scott-Conley, L. (2011). Starting an extensive reading program. The Extensive Reading Foundation document. Available at: <http://extensivereading.net/docs/lois.html>.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: a longitudinal study. *Language Learning*, 48(2), 281-317
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research*, 12, 329-363.
- Schonell, F. J., Meddleton, I. G., & Shaw, B. A. (1956). *A study of the oral vocabulary of adults*. Brisbane: University of Queensland Press.
- Schoot, M., Vasbinder, A. L., Horsley, T. M., & van Lieshout, E.C.D.M. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading*, 31 (2), 203–223.
- Schwanenflugel, P., & Ruston, H. (2008). Becoming a fluent reader: From theory to practice. In M. Kuhn & P. Schwanenflugel (Eds.), *Fluency in the classroom* (pp. 1-16). New York: Guilford Press.
- Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 200–219). Ann Arbor, MI: University of Michigan Press.
- Seow, A. (1999). What do we really want out of USSR? *Teaching of English Language and Literature*, 15(2), 22-24.
- Shanghai English Curriculum Standards for Primary and Junior Middle School* (2004). Shanghai Curriculum and Teaching Materials Reform Commission (eds). Shanghai: Shanghai Education Press.
- Shen, Z. Y. (2004). Yiyi zai shijian mubiao zai chengxiao [Significance and objectives of bilingual education]. *Zhongguo Jiaoyubao*, May 13, 2.
- Shen, M, Y. (2008). EFL Learners' responses to extensive reading: Survey and pedagogical applications. *The Reading Matrix*, 8(2), 111-123.
- Shi, B. H. (1997). Jiaojishi yuyan jiaoxue ershiwu nian [Twenty-five years of communicative language teaching]. *Foreign Language Teaching and Research*, 3, 69-73/83.

- Shi, Z. M. (1994). Huaren shequ xuesheng yuwen yingyong nengli diaocha [An investigation into the students' ability to use Chinese in Chinese community]. *Yuyan jiaoxue yu yanjiu [Language Teaching and Research]*, 1, 105-119.
- Shillaw, J. (1999). An evaluation of a semi-intensive reading programme. On the website of English Curriculum Development. (pp. 79-90). Available at: <http://www.tulips.tsukuba.ac.jp/limedia/dlam/B15/B1511855/8.pdf>.
- Shin, F. (2001). Motivating students with Goosebumps and other popular books. *California School Library Association Journal*, 25(1), 15-19.
- Sin, M. (2007). *An evaluation of the effectiveness of a school-based Chinese extensive reading curriculum for junior secondary students*. Unpublished Ph.D thesis, The University of Hong Kong, Hong Kong.
- Smith, F. (1983). Reading like a writer, *Language Arts*, 60 (5), 558-567.
- Smith, P. J., & Smith, S. N. (1999). Differences between Chinese and Australian students: Some implications for distance educators. *Distance Education*, 20 (1), 64-80.
- Sonbul, S., & Schmitt, N. (2010). Direct teaching of vocabulary after reading: Is it worth the effort? *ELT Journal*, 64, 253-260.
- Spolsky, B. (1989). Conditions for second language learning. Oxford: Oxford University Press.
- Stahl, S. (1999). *Vocabulary development*. Cambridge, MA: Brookline Books.
- Stahl, S. A. (2005). Four problems with teaching word meanings (and what to do to make vocabulary an integral part of instruction). In E. Hiebert & M. Kamil (Eds.), *Teaching and researching vocabulary* (pp. 95-114). Mahwah, NJ: L. Erlbaum.
- Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings*. London: Lawrence Erlbaum Associates Publishers.
- State Education Commission. (1986). *Quanrizhi zhongxue yingyu jiaoxue dagang* [English syllabus for full-time secondary schools]. Beijing: People's Education Press.
- State Education Commission. (1988). *Jiunian yiwu jiaoyu quanrizhi chuji zhongxue yingyu jiaoxue dagang (chushengao)* [English syllabus for nine-year compulsory education full-time junior secondary schools (preliminary draft)]. Beijing: People's Education Press.

- State Education Commission. (1990). *Quanrizhi zhongxue yingyu jiaoxue dagang (xiudingben)*[English syllabus for full-time secondary schools]. Beijing: People's Education Press.
- State Education Commission. (1992). *Jiunian yiwu jiaoyu quanrizhi chuji zhongxue yingyu jiaoxuedagang(shiyong)*[English syllabus for nine-year compulsory education full-time junior secondary schools]. Beijing: People's Education Press.
- State Education Commission. (1993). *Quanrizhi gaoji zhongxue yingyu jiaoxue dagang(chushengao)*[English syllabus for full-time senior secondary schools]. Beijing: People's Education Press.
- State Education Commission. (1996). *Quanrizhi gaoji zhongxue yingyu jiaoxue dagang (gong shiyanyong)* [English syllabus for full-time senior secondary schools]. Beijing: People's Education Press.
- Stern, H. H. (1972). Introduction. In M. Swain (eds.), *Bilingual schooling: Some experiences in Canada and the United States*, (pp. 1-6). Toronto: OISE Press.
- Stone, R. (1999). English in Hong Kong: Word knowledge skills of science undergraduates. *Hong Kong Journal of Applied Linguistics*, 4, 93-100.
- Storey, P., Yu, V., Luk, C. M., Ma, A., Pang, M., Wong, C., et al. (1997). *Feasibility Study on Extensive Reading in Chinese and in English to all levels of Primary and Secondary School*. Hong Kong: Hong Kong Institute of Education.
- Stothard, S. E., & Hulme, C. (1992). Reading comprehension difficulties in children: The role of language comprehension and working memory skills, *Reading & Writing*, 4, 245-256.
- Strong, G. (1996). Using literature for language teaching in ESOL. *Thought Currents in English Literature*, 69, 291-305. ERIC Document Reproduction Service No. ED407860.
- Stuart, K. (1990). Developing extensive reading skills with culturally relevant folktales. *TESL Reporter*, 23(1), 3-4.
- Sun, Y. C. (2003). Extensive reading online: An overview and evaluation. *Journal of Computer Assisted Learning*, 19(4), 438-446.
- Sun, L. (2009). *Exploration on the modes of college English teaching guided by input and output theories*. Unpublished MA dissertation, Jilin Normal University, Changchun, China.

- Susser, B., & Robb, T. N. (1990). EFL extensive reading instruction: Research and procedure. *JALT Journal*, 12(2), 161-185.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125- 144). Oxford: Oxford University Press.
- Swain, M. (2000). French immersion research in Canada: Recent contributions to SLA and applied linguistics. *Annual Review of Applied Linguistics*, 20, 199-212.
- Swain, M., & Johnson, R. K. (1997). Immersion education: A category within bilingual education. In R.K. Johnson and M. Swain (Eds), *Immersion education: International perspectives*, (pp. 1-16). Cambridge: Cambridge University Press.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive process they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.
- Swan, M. (2005). Legislation by hypothesis: The case of the task-based instruction. *Applied Linguistics*, 26 (3), 376-401.
- Tadman, J. (1980). *How to make the most of graded readers on cassette*. Harlow, Essex: Longman.
- Taguchi, E., Takayasu-Maass, M., & Gorsuch, G. (2004). Developing reading fluency in EFL: How assisted repeated reading and extensive reading affect fluency development. *Reading in a Foreign Language*, 16, 70-69.
- Taillefer, G. F. (1996). L2 reading ability: Further insight into the short-circuit hypothesis. *The Modern Language Journal*, 80, 461–477.
- Takase, A. (2003). *The effects of extensive reading on the motivation of Japanese high school students*. Unpublished PhD dissertation, Temple University, Japan.
- Takase, A. (2005). Teachers motivated by students' extensive reading: A case study of teachers' motivation to start reading English books. In K. Bradford-Watts, C. Ikeguchi, & M. Swanson (Eds.) *JALT 2005 Conference Proceedings*. Tokyo. Available at:
<http://jalt-publications.org/archive/proceedings/2005/E084.pdf>.
- Takase, A. (2007a). Japanese high school students' motivation for extensive L2

- reading. *Reading in a Foreign Language*, 19, 1–18.
- Takase, A. (2007b). Extensive reading in the Japanese high school setting. *The Language Teacher*, 31 (5), 7-10.
- Takase, A. (2009). The effects of different types of extensive reading materials on reading amount, attitude and motivation. In A. Cirocki (Ed.), *Extensive reading in English language teaching* (pp. 451-465). Munich, Germany: Lincom.
- Tan, D. X. (1998). *On improving college English students' reading proficiency at Fujian Normal University*. Unpublished Ph.D dissertation, University of South California, USA.
- Tanaka, H., & Stapleton, P. (2007). Increasing reading input in Japanese high school EFL classrooms: An empirical study exploring the efficacy of extensive reading. *The Reading Matrix*, 7(1), 115-131.
- Thompson, I., & Rubin, J. (1996). Can strategy instruction improve listening comprehension? *Foreign Language Annals*, 29 (3), 331-342.
- Thornbury, S. (1999). *How to teach grammar*. Harlow: Longman.
- Thornbury, S. (2002). *How to teach vocabulary*. London: Longman.
- Thorndike, R. L. (1973). Reading as reasoning. *Reading Research Quarterly*, 9 (2), 137-147.
- Tommola, J. (1979). *Some parameters of simplification*. Communication studies paper, MSc in Applied Linguistics. Edinburgh. Dept. of Linguistics, University of Edinburgh, mimeo.
- Tsang, W. (1996). Comparing the effects of reading and writing on writing performance. *Applied Linguistics*, 17, 210-233.
- Tudor, I., & Hafiz, F. (1989). Extensive reading as a means of input to L2 reading. *Journal of Research in Reading*, 12, 164-178.
- Tunmer, W. E., & Hoover, W. A. (1992). Cognitive and linguistic factors in learning to read. In P. B. Gough, L. C. Ehri, and R. Treiman (Eds.), *Reading Acquisition*. Hillsdale, NJ: Erlbaum.
- Twaddell, F. (1980). Vocabulary expansion in the TESOL classroom. In K. Croft (Eds.), *Reading on English as a second language (2nd edition)*, (pp. 439-457). Cambridge: Winthrop.
- Vandergrift, L. (2005). Relationships among motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied*

Linguistics, 26, 70–89.

- Vandergrift, L. (2006). Second language listening: Listening ability or language proficiency. *The Modern Language Journal*, 90, 6-18.
- Vandergrift, L., & Tafaghodtari, M. (2010). Teaching students how to listen does make a difference: An empirical study. *Language Learning*, 60, 470–497.
- Wagner, R. K., Piasta, S. B., & Torgesen, J. K. (2006). Learning to read. In M. J. Traxler & M. A. Gernsbacher (ed), *Handbook of psycholinguistics* (2nd edition) (pp. 1111-1142). Amsterdam: Academic Press.
- Wang, T. J. (2010). Speaking anxiety: More of a function of personality than language achievement. *Chinese Journal of Applied Linguistics*, 33 (5), 95-109.
- Wang, X. C. (2008). *A Study of an English extensive reading experiment in a rural high school--Linzhang high school*. Unpublished MEd thesis, Hebei Normal University, Sijiazhuang, China.
- Waring, R. (1997). Graded and extensive reading: Questions and answers. *The Language Teacher*, 21(5), 9-12.
- Waring, R. (2009). The inescapable case for extensive reading. In A. Cirocki (Ed.), *Extensive reading in English language teaching* (pp. 93-111). Munich, Germany: Lincom.
- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130-163.
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20(2), 232-245.
- Webb, S., & Rodgers, M. P. (2009). Vocabulary demands of television programs. *Language Learning*, 59 (2), 335-366.
- Wendelin, K., & Zinc, R. (1983). How students make book choices. *Reading Horizons*, 23, 84-88.
- West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.
- White, H. (2004). Nursing instructors must also teach reading and study skills. *Reading Improvement*, 41, 38-50.
- Widdowson, H. G. (1976). The authenticity of language data. In J. F. Fanselow, & R. H. Crymes (Eds.), *ON TESOL '76*. Washington: TESOL.

- Widdowson, H. G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.
- Widdowson, H. G. (1979). *Explorations in applied linguistics*. Oxford: Oxford University Press.
- Widdowson, H. G. (1990). *Aspects of language teaching*. Oxford: Oxford University Press.
- Wiesendanger, K. D., & Bader, L. (1989). SSR: Its effect on students' reading habits after they complete the program. *Reading Horizons*, 29 (3), 162-166.
- Wiesendanger, K. D., & Birlem, E. D. (1984). The effectiveness of SSR: An overview of the research. *Reading Horizons*, 24 (3), 197-201.
- Wilkins, D. A. (1976). *Notional syllabuses*. Oxford: Oxford University Press.
- Williams, E. (2007). Extensive reading in Malawi: inadequate implementation or inappropriate innovation? *Journal of Research in Reading*, 30 (1), 59-79.
- Williams, J. (2007). Literacy in the curriculum: Integrating text structure and content area instruction. In D. McNamara (Eds.), *Reading comprehension strategies* (pp. 199-219). New York: L. Erlbaum.
- Williams, R. (1986). Top ten principles for teaching reading. *ELT Journal*, 40, 42-45.
- Wittrock, M. C. (1983). Writing and the teaching of reading. *Language Arts*, 60 (5), 600-606.
- Wodinsky, M., & Nation, I. S. P. (1988). Learning from graded readers. *Reading in a Foreign Language*, 5, 155-161.
- Wu, Y. A. (1998). What do tests of listening comprehension test? – A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing*, 15 (1), 21-44.
- Xu, J. (2005). SPSS in Windows: ANOVA. *EDPR 7/8542*, Spring 2005.
- Yamashita, J. (2007.) The relationship of reading attitudes between L1 and L2: An investigation of adult EFL learners in Japan, *TESOL Quarterly*, 41(1), 81-105.
- Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, 36, 661-672.
- Yang, A. (2001). Reading and the non-academic learner: A mystery solved. *System*, 29, 451-466.
- Yang, J. (2006). Learners and users of English in China. *English Today*, 22 (2),

- Yao, T. (2005). *The application of metacognition to English extensive reading in senior high School*. Unpublished Med thesis, Hebei Normal University, Sijiazhuang, China.
- Yi, B. (2010). *A study on English curriculum change of China's basic education during the thirty years of reform and opening up (1978-2008)*. Unpublished Ph.D thesis. Hunan Normal University, Changsha, China.
- You, C., & Chen, S. (2009). Applying authentic materials to EFL extensive reading in senior high schools in China. In A. Cirocki (Ed.), *Extensive reading in English language teaching* (pp. 311-331). Munich, Germany: Lincom.
- Yu, L. (2001). Communicative language teaching in China: Progress and resistance. *TESOL Quarterly*, 35 (1), 194-198.
- Yu, V. (1995). Organizing an extensive reading book programme. In M. Ingham, & N. Bird (Eds.), *Learning how to learn* (pp. 101-117). Hong Kong: Hong Kong Institute of Education.
- Yu, V. (1999). Promoting second language development and reading habit through an extensive reading scheme. In Y. M. Cheah & S. M. Ng (Eds.), *Language instructional issues in Asian classrooms* (pp. 59-74). Newark, DE: International Development in Asia Committee, International Reading Association.
- Yu, V., & Atkinson, P. A. (1988a). An investigation of the language difficulties experienced by Hong Kong secondary school students in English medium schools: I. The problems. *Journal of Multilingual and Multicultural Development*, 9, 267-284.
- Yu, V., & Atkinson, P. A. (1988b). An investigation of the language difficulties experienced by Hong Kong secondary school students in English medium schools: II. Some casual factors. *Journal of Multilingual and Multicultural Development*, 9, 307-322.
- Yu, V., & Wong, C. (1991). The Hong Kong Extensive Reading Scheme (HKERS)—an extensive reading program for junior secondary school students. *Institute of Language in Education Journal*, 8, 188-192.
- Zhang, H. (2005). *Instructing extensive reading (ER) to college students in China*. Unpublished MEd thesis, Southwest China Normal University, Chongqing,

China.

- Zhang, J. X. (2011). Lun gaozhong xinyingyu kecheng biao zhun dui xuesheng yingyu yuyan nengli de yingxiang [Research on the new round English language curriculum standard on the students' English language proficiency.]. *Theory and Practice of Contemporary Education*, 3(1), 77-79.
- Zhang, S. F. (2012). *Teaching English extensive reading in senior high school*. Unpublished Master of ELT thesis, Central China Normal University, Wuhan, China.
- Zhang, W. J. (2002). Shuangyu jiaoxue de xingzhi tiaojian ji xiangguan wenti [The nature and conditions of bilingual education and some related problems]. *Yuyan Jiaoxue yu Yanjiu*[*Language Teaching and Research*] , 4, 20-6.
- Zhang, Y. H. (2003). *On the feasibility of offering an extensive reading course in senior middle school English classes*. Unpublished MEd thesis, Shan'xi Normal University, Xi'an, China.
- Zhang, Z. F. (2003). Shuangyu jiaoxue de yuanze moshi ji celve [Principles, models and strategies of bilingual education]. *Zhongxiaoxue Jiaoxue Yanjiu* [*Research on Teaching in Primary and Secondary School*], 2, 15-16.
- Zhang, Z. L. (2010). *Kaoshi chengji pingjia de qufendu xindu xiaodu deng gainian* [*The concepts of discrimination, reliability, validity in the evaluation of the test scores*]. Baidu documents, available at: <http://wenku.baidu.com/view/1f10471b6bd97f192279e91f.html>.
- Zhao, W. H. (2006). *Feasible analysis and tactical research on having English extensive reading lesson in high school*. Unpublished MEd thesis, Jiangxi Normal University, Nanchang, China.
- Zhou, B. L., & Chen, J. (2010). The impact of teacher feedback on the long-term improvement in the accuracy of EFL student writing. *Chinese Journal of Applied Linguistics*, 33 (2), 18-34.
- Zhu, H. M. (2003). Globalization and new ELT challenges in China. *English Today*, 19, 36-41.
- Zhu, P. (2003). Lun Shanghai zhongxiaoxue shuangyu jiaoxue shiyan [On experimentation with bilingual education in Shanghai's primary and secondary schools]. *Kecheng, Jiaocai, Jiaofa*, 6, 52-58.
- Zhu, Y. L. (2011). *An application study of learner autonomy in English*

extensive-reading teaching of senior high school. Unpublished MEd thesis, Northwest Normal University, Lanzhou, China.

Zipperer, F. M. J., Worley, M. T., Sission, M. W., & Said, R. W. (2002). Literacy education and reading programs in the secondary School: Status, problems, and solutions. *NASSP Bulletin*, 86 (3), 1-17. Available at: <http://bul.sagepub.com/content/86/632/3>.

Zuo, H. Q. (2002). *Waiyu jiaoxue zhanwang [Outlook of foreign language teaching]*. Shanghai: East China Normal University Press.