Orthographic and phonological processing in English word learning

Rosa Kit Wan Kwok

Doctor of Philosophy

University of York

Psychology

November 2014

Abstract

This thesis investigates the process of orthographic and phonological word learning in adults. Speed of reading aloud is used as the main measure, specifically the reduction in naming reaction times (RTs) to short and long novel words through repetition and the convergence of RTs to short and long items. The first study (Chapter 2) fully described this fundamental learning paradigm and it is then used to compare various types of training in different groups of readers in the following chapters.

Second, the role of phonology in visual word learning was investigated in Chapter 3. Novel words that received the training of both orthography and phonology (reading aloud condition) was found to be more efficient and effective compared to solely training the phonology of the novel words (hear-and-repeat with and without distractors). Yet, all three experiments in Chapter 3 also showed that the establishment of a phonological representation of a novel word can be sufficient of result in representations in the mental lexicon even without any encounter with the orthographic form of the novel word. Linear mixed effect modelling also found that literacy and phonological awareness made a significant contribution to nonwords naming speed when vocabulary and rapid digit naming were taken into account. Expressive vocabulary was found to be a significant predictor of the change in naming speed across the learning session when the effects of literacy, phonological awareness were controlled.

Third, Chapter 4 then involved the repeated presentation of interleaved high-frequency words, low-frequency words and nonwords to native speakers of English in two testing sessions 28 days apart. Theoretical interest lies in the relative effects of length on naming latencies for high-frequency words, low-frequency words and nonwords, the extent to which those latencies (RTs) converge for shorter and longer words and nonwords, and the persistence of training/repetition effects over a 28-day retention interval. Finally, Chapters 5 and 6 try to bring these theories in a more applied context to understand orthographic word learning in adults with dyslexia and in bilingual speakers.

Table of Contents

Abstract……………………………………………………………………………………2

List of Tables…………………………………………………………………….....10

List of Figures……………………………………………………………………….12

Acknowledgments……………………………………………………………………...15

Declaration…………………………………………………………………….........18

1 Chapter 1: Orthographic and phonological word learning in children and adults.... 19

1.1 Introduction 19

1.2 What are the characteristics of successful orthographic learning?..................... 20

1.3 The benefit of artificial word learning.. 22

1.4 Predictors of successful orthographic learning 22

1.4.1 Alphabetic and phonological skills 23

1.4.2 Implicit phonological awareness/fluency skills 25

1.4.3 Vocabulary 27

1.4.4 Working memory 29

1.5 Word learning in children.. 31

1.6 Word learning in adults 34

1.7 Tapping into the process of orthographic learning.. 40

1.8 How can the current computational models explain the mechanisms of word learning?.. 43

1.8.1 The visual feature and letter units 46

1.8.2 The non-lexical route 47

1.8.3 The lexical route.. 47

1.8.4 The transition from serial to parallel processing... 49

1.9 Conclusions and thesis outline 52

2 Chapter 2: Visual word learning in skilled readers of English 53

2.1 Introduction 53
2.2 Experiment 1: learning through repeated exposure 56
  2.2.1 Method 56
  2.2.2 Result 58
  2.2.3 Discussion 61
2.3 Experiment 2: item-specific or general learning? 63
  2.3.1 Method 63
  2.3.2 Result 64
  2.3.3 Discussion 70
2.4 Experiment 3: Long-term retention of new lexical entries 71
  2.4.1 Method 72
  2.4.2 Results 73
  2.4.3 Discussion 79
2.5 General Discussion 79
  2.5.1 The possible contribution of blocking / list context effects 80
  2.5.2 Creation of lexical representations and the modulation of the length effect83
  2.5.3 Length and neighbourhoods 84
  2.5.4 Retention versus decay of lexical representations 86
2.6 Conclusions 88
3 The role of phonology in visual word learning 90
3.1 Introduction 90
  3.1.1 Lexical equality hypothesis 90
  3.1.2 Orthographic learning in children 92
  3.1.3 Orthographic learning in adults 94
  3.1.4 Predictors of word reading in adults 99

3.1.5 Orthographic consistency effect 100

3.2 Experiment 4: The role of phonology in orthographic learning (within subject design) 100

3.2.1 Method 101

3.2.2 Result 107

3.2.3 Predictors of initial nonword reading speed and novel word learning 112

3.2.4 Discussion 115

3.3 Experiment 5: The role of phonology in orthographic learning (between subject design) 118

3.3.1 Method 118

3.3.2 Result 120

3.3.3 Discussion 127

3.4 Experiment 6: Phonological learning with distractors 128

3.4.1 Method 129

3.4.2 Results 135

3.4.3 Discussion 140

3.5 General Discussion 141

3.5.1 Is training the phonology of the word sufficient to build representation in the mental lexicon? 141

3.5.2 Is there automatic and mandatory activation of orthographic codes of English novel words when participants hear them verbally? 143

3.5.3 Predictors of orthographic learning 144

3.5.4 Conclusion 148

4 Reading and lexicalisation in English 150

4.1 Introduction 150

4.1.1 How can the DRC model explain the interaction between lexicality and length 150

4.1.2 Priming effect for real words 152

4.1.3 The present experiment 153

4.2 Experiment 7: The process of reading and lexicalisation in English 153

4.3 Methods 153

4.3.1 Participants 153

4.3.2 Materials 153

4.3.3 Procedure 155

4.4 Result 157

4.4.1 Accuracy 157

4.4.2 Naming latency 157

4.5 Discussion 169

4.5.1 Word learning and the DRC model 170

4.5.2 Lexicalization of new words and priming of words 171

4.5.3 Frequency effect for familiar words are modulated by recent experience 171

4.5.4 Memory retention of learned materials 173

4.5.5 Conclusion and future direction 174

5 Visual word learning in adults with dyslexia 175

5.1 Introduction 175

5.2 Experiment 8: Visual word learning in adults with dyslexia 181

5.2.1 Method 181

5.2.2 Participants 181

5.2.3 Test Battery 181

5.2.4 Stimuli and procedure 183

5.3 Result 184
5.3.1 Performance on the test battery 184
5.3.2 Performance of the experimental task 184
5.4 Discussion 196
6 Visual word learning in Chinese speakers 202
6.1 Introduction 202
6.1.1 Traditional Chinese writing system 203
6.1.2 Second language acquisition 204
6.1.3 How English is taught in Hong Kong 205
6.1.4 Cross-language interaction 207
6.1.5 Hong Kong speakers' phonological awareness ability 210
6.1.6 The current experiment and research question 211
6.2 Methods 211
6.2.1 Participants 211
6.2.2 Materials and procedure 212
6.3 Results 212
6.3.1 Accuracy 212
6.3.2 Naming latency 213
6.4 Discussion 221
6.4.1 The interdependence model in English word learning 222
6.4.2 The phonological core view model in English word learning 223
7 General Discussion 228
7.1 Summary of key findings 228
2.6.1 Chapter 2: Visual word learning in skilled readers of English 228
2.6.2 Chapter 3: The role of phonology of visual word learning 230

2.6.3 Chapter 4: Reading and lexicalisation in English 233

2.6.4 Chapter 5: Visual word learning in adults with dyslexia 235

2.6.5 Chapter 6: Visual word learning in Chinese native speakers.................... 237

7.2 Connectionist model and the Connectionist Dual Process (CDP++) model 239

2.6.6 Connectionist model 239

2.6.7 How words are processed and read... 240

2.6.8 Connectionist Dual Process (CDP++) model... 242

7.3 Alternative explanation (the power law) 244

2.6.9 Are the characteristics of word learning similar to those in general skill acquisition? 244

2.6.10 Can the power law explain the reduction of length effect?....................... 246

7.4 Implications of the findings.. 246

7.5 Future directions.. 248

7.6 Conclusions 252

1 Appendix 1…………………………………………………………………...254

1.1 Analysis of the results in Experiments 1, 2, and 3 of Chapter 2………………..254

1.1.1 Experiment 1 ……………………………………………………………..254

1.1.2 Experiment 2 ……………………………………………………………..255

1.1.3 Experiment 3 …………………………………………………….…………258

2 Appendix 2…………………………………………………………………...262

2.1 Analysis of the results in Experiments 4, 5, and 6 of Chapter 3………………..262

2.1.1 Experiment 4…………………………………………………………262

2.1.2 Experiment 5………....……………………………………………………264

2.1.3 Experiment 6…………………………………………………………268

3 Appendix 3…………………………………………………………….........270

3.1 Analysis of the result in Experiment 7 of Chapter 4……………………….….270

3.1.1 Experiment 7…………………………………………………………270

4 Appendix 4…………………………………………………………………...277

4.1 Analysis of the result in Experiment 8 of Chapter 5……………………………277
  4.1.1 Experiment 8………………………………………………………………277
5 Appendix 5……………………………………………………………………...282
5.1 Analysis of the result in Experiment 9 of Chapter 6……………………………282
  5.1.1 Experiment 9……………………………………………………………....282
6 Appendix: Stimulus of each experiment………………………………………..289
6.1 All the stimulus in each set……………………………………………………..289
6.2 The distribution of sets in each experiment…………………………………….....290
6.3 The stimuli for the orthographic distractors in Experiment 6…………………......293
6.4 Words stimulus used in Experiment 8……………………………………….....296
7 Reference……………………………………………………………………..298

List of Tables

Table 2.1. *Mean and standard deviation of bigram frequency, neighborhood size, reading speed and accuracy of the four and seven-letter nonwords of Set A from the pilot study.* 58

Table 2.2. *Mean latencies of correct, trimmed responses, standard deviation (SD), and per cent correct responses for 4- and 7-letter nonwords in blocks 1 to 10 of day 1 and day 7 in Experiment 1.* 59

Table 2.3. *Mean and standard deviation of bigram frequency, neighborhood size and reading speed of the four and seven-letter nonwords of Set B, C and D from the pilot study.* 65

Table 2.4. *Mean latencies of correct, trimmed responses, standard deviation, and per cent correct responses for 4- and 7-letter trained nonwords in blocks 1 to 10 and for untrained nonwords in blocks 1 and 10 only in Experiment 2.* 66

Table 2.5. *Mean and standard deviation of bigram frequency, neighborhood size, reading speed and accuracy of the four and seven-letter nonwords of Set E from the pilot study.* 74

Table 2.6. *Mean latencies of correct, trimmed responses, standard deviation (S.D.), and per cent correct responses for 4- and 7-letter nonwords in blocks 1 to 10 of day 1 and day 7 in Experiment 3.* 75

Table 3.1. *Mean and standard deviation of bigram frequency, neighbourhood size and reading speed of the four- and seven-letter nonwords from the pilot study.* 103

Table 3.2. *Illustration of the distribution of training and test sessions.* 105

Table 3.3. *Mean latencies of correct, trimmed responses, standard deviation (S.D.), and percent correct responses for 4- and 7-letter nonwords in Session 1 and 2 in Experiment 4.* 110

Table 3.4. *Correlations among the predictor variables, and between the predictor variables and naming RTs for 7-letter nonwords in block 1 of day 1.* 113

Table 3.5. *Illustration of the distribution of training and test sessions in the group that received visual training.* 120

Table 3.6. *Illustration of the distribution of training and test sessions in the group that received phonological training.* 120

Table 3.7. *Mean latencies of correct, trimmed responses, standard deviation (S.D.), and percent correct responses for 4- and 7-letter nonwords in block 1 and 10 in Experiment 5.* 122

Table 3.8. *Mean and standard deviation of bigram frequency, neighbourhood size and reading speed of the seven-letter nonwords. Range of the 7-letter nonwords reading speed is also included.* 132

Table 3.9. *Illustration of the distribution of training and test sessions.* 133

Table 3.10. *Mean latencies of correct, trimmed responses, standard errors (SE), per cent correct responses for 7-letter nonwords and red dot fillers in each condition of Experiment 6.* 138

Table 4.1. *Details of the stimuli used in Experiment 8.* 156

Table 4.2. *Mean RTs (with S.D.) and accuracy (with S.D. and percent of correct) for four- and seven-letter high-frequency words, low-frequency words and nonwords presented on Day 1 and Day 28.* 160

Table 5.1. *Results of the dyslexic and typical readers on the psychological test battery.* 186

Table 5.2. *Mean latencies of correct, trimmed responses, standard deviations (S.D.), and per cent correct responses for 4- and 7-letter nonwords in blocks 1 to 10 of day 1 and day 7 in dyslexics and typical readers.* 187

Table 5.3. Correlations among the predictor variables, and between the predictor variables and naming RTs for 7-letter nonwords in block 1 of day 1. 192

Table 6.1. *Mean latencies of correct, trimmed responses, standard deviations (S.D.), and per cent correct responses for 4- and 7-letter nonwords in blocks 1 to 10 of day 1 and day 7 in Chinese and British readers.* 218

List of Figures

Figure 1.1. Mean naming latency (in milliseconds) for unfamiliar words of 3 - 6 letters as a function of word length (in letters) and training session (taken from Manis 1985)....... 31

Figure 1.2. Schematic representation of the discrete threshold identification task (DTI), the continuous threshold identification task (CTI) for a trial (taken from Salasoo et al, 1985) 36

Figure 1.3. Naming RTs of high-frequency words, low-frequency words and nonwords (taken from Weekes (1997) Figure 1).. 41

Figure 1.4. The mean naming RTs across four repetitions (blocks) in Experiment 1 of Maloney et al's (2009) study.. 43

Figure 1.5. The basic dual route theory of reading aloud (modified from Coltheart et al. 2001) 44

Figure 1.6. The computational architecture of the DRC model (taken from Coltheart et al., 2001).. 45

Figure 1.7. The computational architecture of the DRC model (taken from Dodd, Campbell and Worrall, 1996)... 46

Figure 1.8. Effects of length and lexicality on naming latencies from human readers and DRC model (taken from Coltheart et al., 2001, Figure 10). 50

Figure 1.9. Effects of frequency on reaction latencies from human readers and DRC model (adopted from Coltheart et al., 2001, Figure 8).. 51

Figure 2.1. The accuracy of naming 4- and 7-letter nonwords in Blocks 1 to 10 in the trained and untrained conditions. Error bars show 95% confidence intervals................. 60

Figure 2.2. The naming reaction times (RTs) for 4- and 7-letter nonwords in Blocks 1 to 10. Error bars show 95% confidence intervals.. 61

Figure 2.3. The accuracy of naming 4- and 7-letter nonwords in Blocks 1 to 10 in the trained and untrained conditions. 67

Figure 2.4. The naming reaction times (RTs) for 4- and 7-letter nonwords in Blocks 1 to 10 in trained and untrained conditions. Error bars show 95% confidence intervals......... 68

Figure 2.5. The accuracy of naming 4- and 7-letter nonwords in Blocks 1 to 10 in Days 1 and 7. Error bars show 95% confidence intervals.. 76

Figure 2.6. The naming reaction times (RTs) for 4- and 7-letter nonwords in Blocks 1 to 10 in Days 1 and 7. Error bars show 95% confidence intervals. 77

Figure 3.1. Mean naming latency (in seconds) for common words (C), auditorily learned pseudowords (AV) and visually learned pseudowords (A) (taken from Manis 1985). 92

Figure 3.2. Mean number of pronunciations (10 maximum) recalled over five learning trials by higher and lower ability fifth-grade readers in the spelling seen and spelling not seen conditions. (taken from Rosenthal and Ehri. 2008). 94

Figure 3.3. The naming RTs of the reading aloud group in Maloney et al. (2009). 95

Figure 3.4. The mean response times of the case decision group in Maloney et al. (2009) study. 95

Figure 3.5. The accuracy of naming 4- and 7-letter nonwords in Blocks 1 and 10 in the visual and phonological training conditions. Error bars show 95% confidence intervals. 108

Figure 3.6. The naming reaction time (RTs) for 4- and 7-letter nonwords in Blocks 1 and 10 in visual and phonological training conditions of Experiment 4. Error bars show 95% confidence intervals. 111

Figure 3.7. The accuracy of naming 4- and 7-letter nonwords in Blocks 1 and 10 in the visual and phonological Trained and Baseline conditions. 123

Figure 3.8. The naming reaction time (RTs) for 4- and 7-letter nonwords in Blocks 1 and 10 in visual and phonological training conditions of Experiment 5. Error bars show 95% confidence intervals. 126

Figure 3.9. The mean accuracy of the trained and untrained items in the read aloud, hear-and-repeat, hear-and-repeat with orthographic distractors and hear-and-repeat with non-orthographic distractors conditions. Error bars show 95% confidence intervals. 136

Figure 3.10. The naming reaction time (RTs) for trained and untrained items in the read aloud, hear-and-repeat, hear-and-repeat with orthographic distractors and hear-and-repeat with non-orthographic distractors conditions. Error bars show 95% confidence intervals. *with ortho = hear-and-repeat with orthographic distractors condition, with non = hear-and-repeat with non-orthographic distractors condition. 139

Figure 3.11. The naming reaction time difference (RTs diff) between trained and untrained items in the read aloud, hear-and-repeat, hear-and-repeat with orthographic

distractors and hear-and-repeat with non-orthographic distractors conditions. Error bars show 95% confidence intervals. *with ortho = hear-and-repeat with orthographic distractors condition, with non = hear-and-repeat with non-orthographic distractors condition. 139

Figure 4.1. The accuracy of naming 4- and 7-letter high-frequency words, low-frequency words and nonwords in Blocks 1 to 10 in Days 1 and 28. Error bars show 95% confidence intervals and their absence means that no variance was present in the data. 163

Figure 4.2. The naming latency (RTs) for 4- and 7-letter high-frequency words, low-frequency words and nonwords in Blocks 1 to 10 in Days 1 and 28. 164

Figure 5.1. Naming accuracy to 4- and 7-letter nonwords in controls across two sessions (10 blocks per session). Error bars show 95% CIs. 189

Figure 5.2. Naming accuracy to 4- and 7-letter nonwords in dyslexics across two sessions (10 blocks per session). Error bars show 95% CIs. 190

Figure 5.3. Naming RTs to 4- and 7-letter nonwords in dyslexics and controls across two sessions (10 blocks per session). Error bars show 95% CIs. 191

Figure 6.1. Naming accuracy to 4- and 7-letter nonwords in the British group across two sessions (10 blocks per session). Error bars show 95% CIs. 215

Figure 6.2. Naming accuracy to 4- and 7-letter nonwords in the Chinese group across two sessions (10 blocks per session). Error bars show 95% CIs. 216

Figure 6.3. Naming RTs to 4- and 7-letter nonwords in the British and Chinese group across two sessions (10 blocks per session). Error bars show 95% CIs. 217

Figure 7.1. The general framework of the Connectionist model based on Seidenberg and McClelland (1989) (taken from Plaut et al., 1996). 240

Figure 7.2. Schematic description of the Connectionist Dual Process model (CDP+) (taken from Perry et al., 2007) 243

Acknowledgements

First and foremost, I must thank my supervisor, Professor Andy Ellis, for providing me with exceptional support and guidance throughout my doctoral studies. Thank you for your time and patience over the past four years. Your advice was invaluable in helping me to focus on both the micro- and macro-aspect of my research and eventually transforming it to a coherent thesis. I have been extremely lucky to have a supervisor who cared so much about my research, and who responded to my queries so promptly. Motivational as a professional and inspirational as a human being --- this is what you are.

Second, I would like to extend my gratitude to the other members of my research committee. The guidance of Professor Maggie Snowling had ensure my first three experiments start off with a very solid grounds in understanding word learning in typical adults before I move onto understanding the issues in adults with dyslexia. Professor Gareth Gaskell's meticulous nature helped me to generate experiment with precise and careful designs. His expertise on memory consolidation has also provided me with insightful comments on the memory retention aspect of my thesis. Dr. Meesha Warmington's eye for analysis and practical applications kept my thesis grounded and make my thesis not only related to psycholinguistics but also relevant to developmental psychology. Thank you all for the time and energy that you so willingly devoted to my personal development as a researcher.

The Psycholinguistics Research Group has been a warm, supportive and collaborative environment in which to take my first steps as an independent researcher, and I look forward to continuing to work with this diverse and skilled group of people in the future. In particular, my sincere thanks must go to Professor Sven Mattys and Dorina Strori for helping me to check the analysis of linear mixed effect modelling in this thesis. I would also like to thank my friends and colleagues, in particular to Hannah Witherstone and Faye Smith, for constant encouragement that we give each other starting from the days when we completed the masters degree together. I would like to especially thank Katya Redwood, Shinmin Wang, Tianxiao Yang, Karisha George, Rachel Kowert, Gina

Humphreys, Xin Kang, Elaine Tham, Lauren Welbourne, Myrthe Faber, Alex Reid, Victoria Brattan, Sarah Barbes, Darren Dunning, Rebecca Gilbert, Lily Law, Chiew Kin Yung, and Upasana Nathaniel for making the psychology department such a lovely place to work at. In addition, I have had the good fortune to work with two excellent research assistants: Flora Ip --- thank you for your patience and care for recording the naming accuracy of the participants, and Madelein Crouder, for putting testing participants as a high priority to make sure that we spent the shortest amount of time on testing. Thanks a lot to the international office of University of York to help fund my studies with the 'Overseas Continuation Scholarship'.

I would also like to thank my second family --- the welfare team (CWT) of Goodricke College, for making me learn that we can all make a difference in people's lives. Being a college tutor not only fund my studies financially, but also taught me that psychology is not just about sitting in the office and writing papers. I must thank the provost of Goodricke College, Jane Clarbour, for being such a great role model for me throughout the years. No words can express my gratitude to you. You are the best boss I could ever ask for. Thanks Paul Summers, Sarah Hay, Auriel Hamilton, Jon Fanning, Tom Price, Suzy Dodd, Mihai Cocoru, Steve Eyre, Thomas Martin, Laura Howarth, Ling Ding, Bharati Singh, Megan Maisey and Joshua King for bringing a family-like atmosphere to the workplace.

Getting through the thesis required more than academic support, and I have many people to thank for listening to and, at times, having to tolerate me over the past 4 years. I would like to thank Paul Jervis, Tian Tian, Laura Gettings, Samir Abdoulaye, Karen Chan, Rana Akhundlu, Anja Brown, Betty Hou for their wonderful support. I also want to make a special mention of those friends at home who stood by me when I left, when I came back and all the time in between: Grace Leung, Gladys Chow, Charlotte Chu, Dorothy Yu, Alam Lam, Aquila Ng, and Vivian Chu. I also have to specially thank Mr Simon Chow, for being the inspiration for me to pursue further education in Psychology. He reminded me that working in academia is a privilege and I should do my best to overcome all the difficulties that it takes. I also must thank my great British uncle and aunt, Malcolm Maddock and Mo Maddock, for providing the teas and two pairs of ears whenever they are needed! I would also like to thank Tim New for being so supportive over the past four years.

Most importantly, none of this could have happened without my family. My grandmother, Kwan Siu Chan, offered her encouragement through phone calls every week --- despite my own limited devotion to correspondence. To my parents, Theresa Wong and Nelson Kwok, and my sister, Odilia Kwok --- it would be an understatement to say that you all have always been my tower of strength. Thanks for providing me with a loving and warm family environment when I was busy working. Thank you for your patience, love and laughter. This dissertation stands as a testament to your unconditional love and encouragement.

Declaration

This thesis comprises the candidate's own original work and has not been submitted previously or simultaneously to this or any other University for a degree. All experiments were designed and conducted by the candidate under the supervision of Professor Andy Ellis. Selected aspects of the research described in this thesis have been published and presented elsewhere.

Publications

Kwok, R.K.W. & Ellis, A.W. (2014). Visual word learning in adults with dyslexia. *Frontiers in Human Neuroscience*, *8*, 1-12.

Kwok, R.K.W. & Ellis, A.W. (in press). Visual word learning in skilled readers of English. The Quarterly Journal of Experimental Psychology.

Presentations

Ellis, A.W., Kwok, R.K.W., Kaestner, M., Gerin, M., Avdyli, R., Bermúdez, B., & Cuetos, F. (July, 2014). Word learning, grain size and orthographic transparency. Paper presented at the United Kingdom Orthographic Group meeting, Manchester, UK.

Kwok, R.K.W. & Ellis, A.W. (by invitation 2013, July). Orthographic and phonological processing in English word learning. Psycholinguistics Research Group, University of York, UK.

Kwok, R.K.W. & Ellis, A.W. (2013, September). Orthographic processing in English word learning. Paper presented at the Joint Annual Conference of the British Psychological Society Cognitive and Developmental sections, Reading, UK.

Kwok, R.K.W. & Ellis, A.W. (2013, July). Visual word learning in adults with dyslexia. Paper presented at the Twentieth Annual Meeting of the Society for the Scientific Study of Reading, Hong Kong, China.

Kwok, R.K.W. & Ellis, A.W. (2013, March). Orthographic and Phonological Processing in English word learning. Poster session presented at the 11st International Symposium of Psycholinguistics, Tenerife, Spain.

1 Chapter 1: Orthographic and phonological word learning in children and adults

1.1 Introduction

This thesis investigates the process of visual word learning in adults. Speed of reading aloud is used as the main measure, specifically the reduction in naming reaction times (RTs) to short and long novel words through repetition and the convergence of RTs to short and long items. This fundamental learning paradigm is then used to compare various types of training in different groups of readers. Existing research which is relevant to this topic is reviewed in this chapter.

Learning new vocabulary is a life-long endeavour. Young children typically produce their first words at the age of about 12 months. Their vocabulary then grows rapidly until an average young adult has a vocabulary of at least 20,000 words while an educated adult may know 70,000 words or more (Bloom, 2000; Mayor & Plunkett, 2010; McMurray, Horst, & Samuelson, 2012; Oldfield, 1966). Early words are learned entirely in spoken form, but when children learn to read at the age of 5 or 6 years, new words are often learned through the medium of written language or simultaneously in speech and writing (Ehri, 2005). However, for adults, unfamiliar words encountered in text are often words that are unfamiliar in both speech and print. As adults have to learn a lot of new words while learning a new subject or foreign language. Therefore, studying people learning new words is of practical importance.

A lot of work has demonstrated that the recognition of letter patterns (Bowey & Hansen, 1994; Ehri, 1998) and direct connections between the written and spoken forms of words (Ehri, 1992, 1998; Share, 1995, 1999, 2004; Stanovich, 1993) are indispensable for effective and accurate reading. Yet, despite the substantive research designed to understand the interaction among orthographic, phonological and semantic learning (Perfetti & Hart, 2002), it is not particularly clear how an unfamiliar word proceeds to become familiar in the mental lexicon.

The present thesis is driven by the underlying research question: how do unfamiliar words build representation in the mental lexicon? This question can be segregated into four main themes: 1) Acquisition: how much of the learning is orthographic and how much of it is phonological, 2) Unitization: how many exposures are required before a new word can build lexical entries in the mental lexicon and be processed in a unitized way? 3) Mechanism: how can existing computational models in visual word recognition accommodate the process of word learning, and (4) Retention: how long can we retain learned information about novel words?

To develop these questions further, the following literature review will firstly discuss what are the successful features of orthographic word learning (in section 1.2). The benefits of adopting an artificial learning paradigm to address the factors of word learning will be discussed in section 1.3. The predictors of successful orthographic learning will then be examined in section 1.4. The relevant historical and theoretical contexts relating to orthographic and phonological word learning in children and adults will be considered in sections 1.5 and 1.6, while section 1.7 will illustrate the process of how a letter string becomes part of a lexicon. Section 1.8 will discuss how computational models can illustrate the multiple processing levels of visual word recognition. Finally, section 1.9 will provide the framework of this thesis.

1.2 What are the characteristics of successful orthographic learning?

A clear characterization of the outcome of the orthographic learning process is required in order to understand how learning occurs. Treisman (1961) suggested that each individual holds a mental ‘dictionary’ (lexicon), storing representations of all known words (see section 1.8). Perfetti and Hart (2002) extended this idea to the lexical quality hypothesis. It involves having developed *fully specified*, rather than partially specified, internal representations. This means that the input code is sufficient to uniquely identify the word to be read, without the necessity for discriminating between competing, partially activated entries. This represents the *autonomy* in the word recognition process. The idea is that reading skill is supported by knowledge of words, including the reader’s representations of orthography, phonology, morphology and meaning. Perfetti and Hart suggested that a good quality representation is operationalized as the efficient and accurate retrieval of a word's pronunciation, meaning, and/or spelling in response to one of the other constituents.

The lexical quality hypothesis suggests that phonological and orthographic representations are inextricably linked in both directions for familiar words. Indeed there is much evidence to suggest that a word's orthography affects the speed and accuracy of processing its phonology. An example is the orthographic consistency effect: participants find it easier to perform phoneme deletions on items where there is a direct correspondence between letters and target sounds than where there is not. This effect was found in English (Castles, Holmes, Neath, & Kinoshita, 2003), French (Pattamadilok, Perre, Dufau, & Ziegler, 2009; Petrova, Gaskell, & Ferrand, 2011; Ziegler, Petrova, & Ferrand, 2008), and Portuguese (Ventura, Morais, Pattamadilok, & Kolinsky, 2004). The reverse effect of the orthographic consistency effect was also found in English (Perfetti, Bell, & Delaney, 1988), German (Ziegler, Perry, Jacobs, & Braun, 2001) and in a wide range of tasks in including phoneme deletion, spelling, lexical decision, semantic categorization, rime detection, naming and masking tasks. However, such a consistency effect was not always found for pseudowords. Bürki, Spinelli, and Gaskell (2012), Pattamadilok et al., (2007, 2009) showed clear consistency effects for pseudowords, but Ziegler and Ferrand (1998) and Ventura et al. (2004) did not. This may be related to the task as Bürki et al. (2012) employed the lexical decision task whereas Ziegler and Ferrand (1998) used priming task. The issue of whether there is an orthographic consistency effect for English peudowords will be addressed in Chapter 3.

The characteristics of a sophisticated orthographic recognition system should be reached by a child at some certain point in time. Yet, it is plausible, that this progresses in an *item-based* manner (Share, 1995, 1999, 2004). This means that at a certain time, one may be reading some words effortfully and slowly, relying heavily on the context and alphabetic decoding, while other words can be processed automatically and rapidly. Thus, it leads researchers to look for evidence of the existence of full autonomy, specificity, and unconsciousness at the item or word level. If that is the case, what is good way to simulate the development of naturalistic word learning?

1.3 The benefit of artificial word learning

Recent research has shown that adults can learn spoken novel words and affixes in laboratory situations and that given a period of overnight consolidation (Dumay & Gaskell, 2007; Sio, Monaghan, & Ormerod, 2013), these novel words come to behave like known words in psycholinguistic tasks (Bowers, Davis, & Hanley, 2005; Gaskell & Dumay, 2003; Magnuson, Tanenhaus, Aslin, & Dahan, 2003; Merkx, Rastle, & Davis, 2011). The benefit of investigating orthographic effects in spoken language in the laboratory lies in the exquisite methodological control it offers, making it possible to dispense with the between item designs that have characterized much of the work on speech perception (e.g. the natural correlation of frequency effect, age of acquisition and imageability), and instead select a single set of spoken targets whose stimuli characteristics can be manipulated across participants.. This can help to reduce concerns about other uncontrolled factors influencing the results of a study. Nonwords are essentially words being seen for the first time: every familiar word starts out as an unfamiliar word. Thus, it will be useful to investigate how people learn new words by simulating natural word learning in the laboratory. Given a good methodology to understand the process of visual word learning, the question is what factors are important for successful orthographic learning?

1.4 Predictors of successful orthographic learning

The next step forward to understand the process of how orthographic learning develops is to pinpoint the main predictors of this skill. What factors appear to be strongly related to skilled, word-level reading of the form described above? Attaining such predictors has confirmed to be difficult as these predictors often have a strong inter-relationship (Griffiths & Snowling, 2002; Zeegers, 2004). The predictors that have a significant contribution in children's word learning may be different from those for adult word learning as adults often have a large vocabulary size before a new word adds to the mental lexicon. In this section, the reputed predictors of successful orthographic learning in children will be briefly mentioned. The main focus will be on the predictors of adult word learning, whose strengths and limitations with regards to each other will be elucidated.

1.4.1 Alphabetic and phonological skills

It is useful to draw a distinction between implicit and explicit phonological processing when considering the relationship between phonological skills and word recognition (Snowling & Hulme, 1994; Wagner & Torgesen, 1987). Explicit phonological awareness is a metalinguistic ability that requires reflection on, and often the manipulation of, the phonological components of spoken words (Gombert, 1992). In contrast, implicit phonological awareness can be defined as a cognitive process that involve speech codes but without conscious awareness. This section focuses on explicit phonological awareness: implicit phonological awareness will be discussed in section 1.4.2.

Bradley and Bryant (1983) showed that a measure of rhyme ability in young children was a good predictor of their subsequent progress in learning to read. Later studies that show alphabetic decoding is known to account for a large variance in children's word recognition (Adams, 1990; Wagner & Torgesen, 1987), with some estimates of the correlation between nonword reading and word reading being as high as .90 (Firth, 1972). This correlation is borne out in longitudinal studies, which indicate that early alphabetic skills are predictive of later word recognition skills (Muter, Hulme, Snowling, & Stevenson, 2004). In a two-year longitudinal study of 90 British children, Muter et al. (2004) showed the ability of children's phonological awareness (phoneme completion, beginning phoneme deletion, and ending phoneme deletion) and letter-sound-knowledge predict later word recognition skills. Furthermore, children with dyslexia, who have demonstrably poor word-level reading skills, often show a nonword reading deficit (Rack, Snowling, & Olson, 1992; Snowling, 2000). Intervention studies have also shown that phonological awareness may be causally implicated in reading development: in line with the phonological linkage hypothesis, Hatcher, Hulme, and Ellis (1994) demonstrated the combined training of letter knowledge and phonological awareness showed a larger improvement in reading skills than did the other groups who were given equal amounts of teaching concentrated solely on reading or on phonological training.

Bowyer-Crane et al. (2008) compared the efficacy of two randomly assigned interventions for children with weak oral language skills at school entry. One group of children received an intervention promoting phoneme awareness and letter-sound knowledge (P + R), along with practice in guided reading of simple books with a teaching assistant; the other group received a contrasting program targeting oral language skills (OL group training vocabulary, grammar and narrative skills). At the end of 20 weeks of intervention, the P + R group was ahead of the OL group in phoneme segmentation and blending, letter-sound knowledge, and measures of reading and spelling.

Young et al. (2002) is one of the few studies that have investigated the long-term academic consequences of childhood language impairment. A group of children (n = 229), first identified as having speech and/or language impairment in a community-based, longitudinal study at 5 years of age and matched controls, were re-examined during early adulthood (age 19). The children were separated into four groups, including speech impaired only group, language impaired only, speech and language impaired, and control. A comprehensive battery of speech and language, cognitive and achievement tests were completed by subjects. Phonological awareness was found to be a significant unique contributor of spelling achievement in all groups, over and above non-verbal IQ and rapid digit naming. This demonstrates that phonological awareness deficits persist well into adulthood.

However, some studies have shown that the skills important for reading may change as a child grows. Using path analysis, Vellutino, Tunmer, Jaccard, and Chen (2007) suggested that the phonological and decoding skills were found to be stronger and statistically more stable in a Younger group (grade 2 & 3) than in an Older group (grade 6 & 7), whereas the relationship between language comprehension skills (listening comprehension --- the ability to comprehend narrative text presented orally) and reading comprehension tended to be stronger in the Older than in the Young group. This illustrates that language comprehension rather than decoding becomes the dominant process in reading comprehension when the reader has acquired enough facility in word identification to decode in written text.

As several researchers have noted, substantial variance in word reading remains unaccounted for when both alphabetic and phonological awareness skills are taken into account (Nation & Snowling, 2004). This leads to the view that these abilities may be necessary, but not uniquely sufficient, for the development of skilled word recognition. Therefore, it would seem that the transition to skilled orthographic reading, characterized by full specificity and autonomy, may be affected by other factors.

1.4.2 Implicit phonological awareness/fluency skills

In order to understand the distinction between implicit and explicit phonological processes, Wolf and her colleagues proposed the double deficit hypothesis (Wolf, 1997; Wolf & Bowers, 1999). According to this hypothesis, 'phonological deficits and the processes underlying naming speed are separable sources of reading dysfunction' (Wolf & Bowers, 1999, p. 416). Convergent evidence over the last 2 decades has demonstrated that the majority of children with reading difficulties and dyslexia across all language and ages tested have naming-speed deficits (Bowers, Steffy, & Tate, 1988; McBride-Chang & Manis, 1996; Moll, Hulme, Nag, & Snowling, 2013; Wood & Felton, 1994). The naming-speed deficits and the well-known phonological deficits represent two independent sources of word recognition failure whose co-occurrence leads to serious reading difficulty.

The association between visual word recognition and the process supporting naming speed is complex. As discussed by Wolf, Bowers, and Biddle (2000), naming speed is the end product of a combination of both lower level perceptual, attentional, articulatory, and lexical retrieval processes and higher level cognitive and linguistic processes, each of which requires rapid rates of processing. This is particularly the case for numeric stimuli which reach automatic levels of processing. The authors proposed that many of these same processes are also utilized in word recognition processes in reading. In light of this argument, Wagner, Torgesen, Laughon, Simmons, and Rashotte (1993) found rapid naming loaded on a separate factor to phonological awareness and short-term memory task in a confirmatory factor analysis data from children. In line with this, Levy, Abello, and Lysynchuk (1997) found rapid automatized naming (RAN) was a significant predictor of text-reading speed in regression analysis of grade 4 children.

Most of the existing literature interpreted the concurrent link between rapid digit naming and reading rate simply because they share *genera*l demands of rapid execution. That is, the visual stimuli in the task (typically letters, digits, pictures) have to be mapped rapidly to their names, and that these mappings are in a sense ‘arbitrary’. For instance, seeing the digit ‘8’ does not provide the participant with the phonological information needed to say the word ‘eight’. Yet, this explanation was challenged by Savage et al. (2005) who sought to explore the specificity of the association between the two by separating rapid naming into rapid digit and picture naming. The study included 67 children, the majority of whom had very poor reading skills. Regression analysis revealed that the significant predictor of reading rate, which is based on the number of words read per minute, was digit naming speed rather than picture naming speed. Even after further controlling reading accuracy, digit naming was a significant predictor of reading rate whereas phonological awareness tasks predicted reading accuracy and comprehension, which was based on the number of questions answered correctly. This result is in line with other studies indicating that RAN and phonological processing predict different broad components of reading ability (Pennington & Lefly, 2001; Young & Bowers, 1995). The fundamental questions of *why* rapid naming for digits but not for pictures is a predictor of reading rate is yet to be fully answered in the wider RAN literature. Savage et al. (2005) suggested that one possibility is that the difference between the rapid naming of numeric and non-numeric stimuli might reflect differences in sub-lexical processes required for the execution. Picture naming, unlike letter naming, probably requires mandatory access to semantic information (Humphreys, Riddoch, & Quinlan, 1988). Semrud-Clikeman, Guy, Griffin, and Hynd (2000) suggested that it may be the case that picture naming may tap into attentional resources in a way that digit naming does not in older children where letter and number recall have become automatized.

Research into the predictive association between rapid naming and reading has yielded mixed results. In a longitudinal study following young beginner readers, Wagner et al. (1997) found significant distinct contributions of rapid naming and phonological awareness to later word-reading ability. Nevertheless, when prior reading skills were controlled for rapid naming did not account for any unique variance in reading whereas phonological awareness did.

Young et al. (2002) showed that RAN remained to make a unique contribution over phonological awareness and non-verbal IQ to word identification to all adults (including language impaired (LI) and control group), pseudo-word reading for the LI group and to spelling for the non-LI group. This illustrated that while phonological awareness is a robust predictor of reading skill in adulthood, RAN is more specifically relevant to the sub-skills involved in single word reading. The current literature on rapid digit naming has relied heavily on the word learning skills in children. This thesis aims to expand by understand the association of rapid digit naming and word recognition in adults with vocabulary and phonological awareness skills taken into account in Chapter 3.

1.4.3 Vocabulary

Ouellette (2006) drew a practical distinction between the *breadth* of participants' vocabulary and the *depth* of their vocabulary knowledge. Ouelette (2006) suggested that an assessment of vocabulary depth is word definitions where participants provided an oral definition for a set of words --- this is also known to tap into expressive vocabulary. On the other hand, vocabulary breath can be assessed by participants selecting the appropriate pictures to match spoken words --- this is also known as receptive vocabulary. This distinction branches from theoretical work in psycholinguistics that the lexicon is a store of phonological word forms that are independent from, but heavily connected to, semantic representations (Levelt, Roelofs, & Meyer, 1999). Ouellette (2006) found that the breadth and depth of vocabulary showed a differential relationship with different aspects of reading, with depth related to reading comprehension (c.f. Braze, Tabor Shankweiler, & Mencl, 2007) while vocabulary breadth was related to nonword reading. Though the segregation between vocabulary breadth and depth is logical and theoretically motivated, it is blurred by a number of factors being confounded. For example, providing the definitions of words is an assessment of depth of knowledge, but this is not necessarily independent of vocabulary breadth (Nation & Cocksey, 2009). This substantial overlap limits the conclusions that can be drawn from Ouellette's (2006) finding concerning the role of vocabulary knowledge in word reading development.

Though the way in which expressive and receptive vocabulary should be classified is unclear, there is a remarkable stability between early vocabulary knowledge and later school performance. Nation and Snowling (2004) found that vocabulary knowledge accounted for unique variance in children's word reading measured concurrently at 8 years of age and longitudinally when the children's reading was retested 5 years later at 13 years of age, even after decoding (nonword reading) and phonological skills were taken into account.

Scarborough's (1998) meta-analysis study showed kindergarten vocabulary skills to be associated consistently with later reading performance. The median *r* for studies investigating the association between receptive and expressive vocabulary in kindergarten and later reading achievements was .38 (20 samples) and .49 (5 samples), respectively. Out of 19 predictors studied by Scarborough, expressive vocabulary was the significant predictor of later reading after alphabet knowledge, print exposure and story recall were taken into account.

The cognitive basis of word learning differences has not been nearly as well studied in young adults as in learners during the primary school years. Braze et al. (2007) recruited 44 adult participants (age 16 to 24) to understand whether vocabulary captured a unique variance in reading comprehension. As they hypothesized, orally assessed vocabulary knowledge had a unique variance in predicting reading comprehension even after listening comprehension and decoding skill were accounted for.

Acknowledging the fact that vocabulary skills consistently predict reading skills and word learning, Wise, Sevcik, Morris, Lovett, and Wolf (2007) tried to understand the basis of this association. They speculated that vocabulary knowledge may aid in word identification through two routes. The first route may reflect a link between stored phonological representations and specific orthographic patterns. Thus, students with smaller vocabulary size may have difficulty in a word recognition task as they do not have well-established, internalized phonological representations of words to map onto written words. The second route involves depth of vocabulary knowledge and may reflect greater speed in encoding, organizing and retrieving of phonological representations of words.

1.4.4 Working memory

The concept of working memory was originally developed by Baddeley and Hitch (1974) and extended by Baddeley (2000). The working memory model includes a central executive linked directly with three other subsystems: the phonological loop, the visuospatial sketchpad, and the episodic buffer. The central executive is a flexible system responsible for the control and regulation of cognitive processes including temporary activation of long-term memory (Baddeley, 1998) and the coordination of multiple tasks (Baddeley, Della Sala, Gray, Papagno, & Spinnler, 1997). The central executive is underpinned by two systems: the verbal storage system (i.e., the phonological loop; Baddeley, 1986) and the visuospatial sketchpad which is specialized for the processing of material that can be represented in terms of its visual or spatial characteristics (Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999).

A variety of evidence that has expanded in recent years suggests that the process of vocabulary acquisition and verbal short-term memory may be connected. In children, reliable correlations have been obtained between digit span, nonword repetition ability, and vocabulary scores, even when other possible factors such as nonverbal intelligence and age have been taken into account (Gathercole & Baddeley, 1989; Gathercole, Service, Hitch, Adams, & Martin, 1999). Nonword repetition ability, a task that required participants to repeat each nonword accurately immediately after it has been presented, has been shown to be associated with more rapid learning of the phonology of new words by children in experimental tasks (Michas & Henry, 1994).

Studies of word learning in adults also support the view that verbal short-term memory is engaged in phonological learning of new words. Using a nonword learning paradigm, Gupta (2003) presented participants with nonword-picture pairs in which the nonwords were presented auditorially and represented the names of the pictured objects (imaginary animals). There was a significant partial correlation between digit span and word-learning score that was measured by a cued recall task. This study obtained a similar result as Atkins and Baddeley (1998) that nonword repetition ability in adults is highly associated with the rate of learning novel phonological forms that do not closely resemble familiar native words.

Results from a key study by Papagno and Vallar (1995) demonstrate that this association extends to exceptionally strong as well as weak word learning abilities. They compared the nonword repetition and novel word learning abilities of young adults classified as either polyglots (people who were skilful at a minimum of three languages, and were learning a foreign language at university) or non-polyglots. Two main findings were shown. Firstly, the polyglots had remarkably high nonword repetition scores compared to the nonpolyglots. Secondly, nonword repetition was specifically and highly associated to the ability to learn novel words in the word learning task. Combining both results, these findings indicate that the word learning mechanism tapped by nonword repetition activates across the life span, though its operation under some conditions may be supported by the proficient foundation of the user's language. Yet, this study did not address whether the superior nonword repetition ability is a cause or effect of the polyglots' general language skills. This question can only be answered with future longitudinal studies.

The evidence addressed so far draws on findings from both children and adults data. Baddeley, Gathercole, and Papagno (1998) suggested that this evidence indicated that nonword repetition, which taps into the phonological loop component of the working memory model, is significantly constrained by phonological storage capacity, and that this capacity plays a dominant role in supporting learning of the sound structure of new words during vocabulary acquisition. In line with Brown and Hulme (1996), Gathercole (2006) proposed that initial encounters with the phonological forms of novel words are represented in the short-term store, and that these representations form the foundation for a gradual process of building a stable and refined representation of the sound structure across repeated presentations. Thus, based on this proposal, if the participant has a weak verbal short-term memory, the quality of the temporary phonological representation in the phonological loop will be compromised which will result in slower rate of learning.

1.5 Word learning in children

Research shows a relatively small number of exposures (4 – 6 times) appear to be sufficient for acquiring orthographic representations for young children (Manis, 1985; Reitsma, 1983). Manis (1985) taught fifth- and sixth-grade normal and disabled readers to learn the meaning and pronunciation of English unfamiliar words varying in word length and in letter-sound regularity and complexity. By the end of the third session, children had been exposed to each word 10 times, including in counting training and experimental trials on the naming tasks. As shown in Figure 1.1 word length effect remained large for disabled readers in the third test session which suggest that they tended to process words in terms of individual components such as letter patterns, even after considerable practice at recognizing the words. In contrast, normal readers showed a decrease in the size of the word length effect in session 2 (the point when they encountered the unfamiliar words for 5 times), which is consistent with a change from component processing to the processing of words as single units. Reitsma (1983) found a similar effect that 4 exposures of the novel words in reading aloud training were sufficient for Dutch children to retain information about sight words in memory.

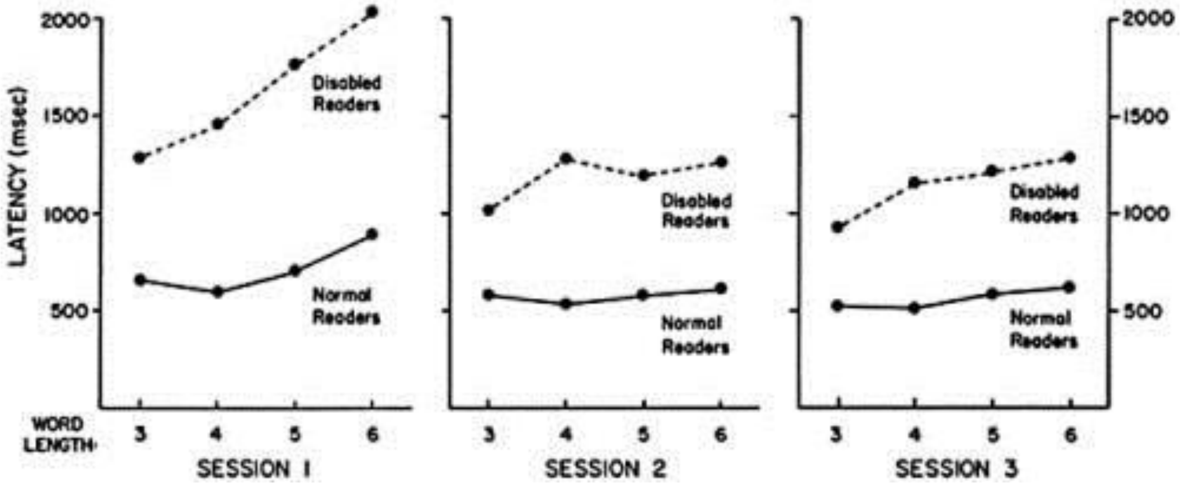

***Figure 1.1*. Mean naming latency (in milliseconds) for unfamiliar words of 3 - 6 letters as a function of word length (in letters) and training session (taken from Manis 1985).**

Share (1995) considered phonological recoding to be the *sine qua non* for the rapid and successful reading acquisition as it forms the foundation of a self-teaching device. Extending an experimental paradigm developed by Reitsma (1983), Share (1999) investigated second-graders' orthographic learning from reading aloud short stories, printed in pointed Hebrew script, with each story containing either four or six repetitions of a nonword that denoted a fictitious object, animal, or place. They read independently, with no guidance or feedback from the experimenter. Three days later, Share tested whether orthographic learning had taken place. An orthographic choice task was used. Each target word was presented alongside a homophone foil (an example in English would be the target word *yait* would be presented alongside the homophone *yate*) and two nonhomophonic foils that shared letters with the target item. Children chose the target on 70% of occasions, five times more often than they chose the homophonic foil. Children also named target items faster than homophone foils, and they were more likely to use the target word pattern, rather than the homophonic spelling pattern, when asked to spell the target words. As there was no difference in learning after either four or six exposures, leading Share to conclude that 8-year-old children show substantial orthographic learning after as few as four exposures to a novel word. This result was replicated and extended by Share (2004) who showed that newly-acquired orthographic information was retained one month later.

The self-learning device theory of Share (1995, 1999, 2004) consists of two main mechanisms. First, basic letter-sound knowledge and decoding skills provide young children with a way of mapping a printed word into its spoken form. This goes in line with the full-alphabetic phase that was suggested by the stage theory (Ehri, 2005) that beginning readers must establish a system of mappings or correspondences between the letters or graphemes of written words and the phonemes of spoken words. Second, this fairly effortful decoding process provides an opportunity to acquire word-specific orthographic information that is needed to gain efficient word recognition.

Using Share's paradigm, Cunningham et al. (2002, 2006) examined this issue in second-and first-grade children learning to read English. Each target novel word appeared six times in a story. Consistent with the difficulty of phonological decoding in English, decoding accuracy was lower than in Share's Hebrew experiments (74% versus upward of 90%), yet, orthography learning occurred. Three days after exposure, children were quicker and more accurate at naming, producing and identifying target words relative to homophonic control words.

Recognizing that children only need a few occurrences of novel words in order to learn them, Kyte and Johnson (2006) try to tease apart the mechanism that supports this rapid word learning process. They re-assessed whether the phonological recoding is a self-teaching mechanism that results in orthographic learning of printed words in English. During a learning phase on Day 1 of testing, the participants performed lexical decisions to real words and pseudo-words under two contrasting experimental conditions - a read aloud condition designed to promote phonological recoding in which items were named prior to lexical decision, and a concurrent articulation condition, designed to attenuate phonological recoding while allowing orthographic processing to occur (participant saying 'LA' from the onset of presentation of novel words). Orthographic learning was evaluated 1 day later with orthographic choice, spelling and naming tasks. Pseudowords learned in the read aloud condition yielded greater orthographic learning on post-test than pseudowords learned with concurrent articulation. Similar conclusion were found by Bowey and Muller (2005) in third graders and De Jong , Bitter, Setten and Marinus (2009) in second graders.

Knowing that phonological recoding is an important process in word learning, Ricketts, Bishop, and Nation (2009) investigated the integration of orthography and phonology by exploring whether exposure to orthography facilitates oral vocabulary learning. Children were trained to associate novel phonological forms with pictures of novel objects. Pictures were used as referents to represent novel word meaning. For half of the nonwords, children were additionally exposed to orthography, although they were not alerted to its presence, nor were they instructed to use it. By the end of training, children had been exposed to each item six times. After the training phase, a nonword-picture matching post-test was used to assess learning of nonword meaning, and a spelling post-test was used to assess learning of nonword orthography. Child showed robust learning for novel spelling patterns after incidental exposure to orthography. Furthermore, there was stronger learning for nonword-referent pairings trained with orthography. Similar result showing that phonology facilities orthographic learning were found by Hu (2008) in children who learn English as a second language and by Duff and Hulme (2012) in 6-year-old British children.

Not only do children learn new words rapidly, but evidence has shown that once the representation of the learned material is built, the memory is retained for a good period of time. Hogaboam and Perfetti (1978) found that training fourth grade (9- to 10-year-old) children on the spoken and written forms of novel words (nonwords) over a period of three days led to faster reading of the same items 10 weeks later. The evidence suggests, therefore, that lexical representations created as the results of a relatively few exposures to novel words can be surprisingly resilient. Yet, this suggestion was criticized by Share (2004) as Nagy and Merman (1987) estimated that children are exposed to millions of printed words each year, which means even rare words would be appearing often enough to refresh diminishing representations. Thus, further research is required to understand how long newly acquired orthographic information is retained.

A similar result has been observed by Martin-Chang, Levy, and O’Neil (2007) in younger children. Extending Archer and Bryant (2001) study, Martin-Chang et al. (2007) taught second grade children novel words in two conditions: context training presented words in stories, and isolated word training presented words on flashcards. The study showed that context training promoted word acquisition beyond the experience from reading words in isolation as children identified approximately 7% more items when the words were presented in a new story context than when the words were presented on flashcards. However, memory performance for words trained in context and in isolation did not differ; children demonstrated excellent retention that reached ceiling effect over an 8-day interval in both conditions.

1.6 Word learning in adults

As mentioned earlier, there are differences in the learning situation for adults acquiring words in reading compared with how children learn new orthographic representations. Adult speakers may sometimes be required to learn new sets of vocabulary. This arises either in the context of mastering a new content area, such as when a student majors in Finance, or in the context of second language learning. These two types of word learning differ in many respects. Most notably, in the former case, both a new concept and its associated label must be learned (e.g. a type of security that signifies ownership in a corporation and represents a claim on part of the corporation's assets and earnings is called 'stock'), but in the latter case, a new label must often be associated with an already familiar concept (e.g. 'argent' is the French translation for the concept 'money'). Normally, both the spoken and written form must be learned. As there are fundamental differences between the way children and adults learn new words, the methodologies that are utilized in the word learning literature in adults are different from those in children.

Salasoo, Shiffrin, and Feustel (1985) trained ten participants in two conditions, one of which participants saw brief presentations of whole target item followed by a mask (the discrete threshold identification, DTI). The other condition consisted of a series of DTI display configurations presented in a very rapid succession with the duration of the stimulus item relative to the mask increasing by a small amount with presentation (the continuous threshold identification, CTI). Figure 1.2 showed the schematic representation of the DTI and CTI conditions. The last item and mask in a trial was immediately followed by the appearance of a small question mark in the centre of the screen, signalling the identification phase of the trial. When the question mark appeared, the subjects attempted to identify the item that had been presented by saying it aloud. After the subjects made their response, the target item that had been presented would be shown on the screen, this allowed the subjects to score their accuracy by pushing the appropriate button on the keypad. The participants were told that their responses were being recorded and that the recordings would be checked for accuracy at a later time. Yet, the verbal responses were not in fact recorded. Participants were trained in the 10 sessions over 12 days in which words and pseudowords were presented 30 times. In each session, an equal number of DTI and CTI trials was presented in a mixed list composed of half words and half pseudowords.

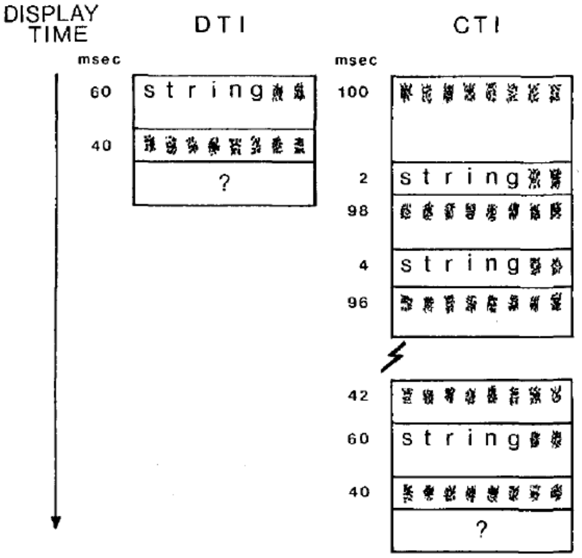

Figure 1.2. **Schematic representation of the discrete threshold identification task (DTI), the continuous threshold identification task (CTI) for a trial (taken from Salasoo et al, 1985)**

Results illustrated that by approximately the sixth presentation, word and pseudo words were identified equally accurately, suggesting that the learning of novel orthographic forms is rapid in adults, as it is in children. Eight of the ten participants were re-tested a year later with a mixed list of new and old words and pseudo words (learned pseudo words were classified as old pseudo words). Participants performed in two 90-min sessions on consecutive days. Results showed that performance improved between the two experimental sessions, suggesting the presence of a warm-up effect). As the pattern of results was similar on the two days, individual subject data for each experimental condition were collapsed across sessions. In both DTI and CTI conditions, old words and old pseudo words were identified equally accurately. Salasoo et al (1985) interpreted the result as the learning had been completed for the old pseudowords and its representational codes in the mental lexicon were still accessible across a gap of 12 months during which the learned pseudowords would not have been experienced. No differences were observed between performance on new and old words. The difference between new and old pseudowords had begun to decrease by the third presentation. Though this result of the study was very informative, this study was flawed in two ways. Firstly, the result was limited to the eight participants who returned for the follow-up session after 12 months. Secondly, the result relied on the participants' self-monitoring response which may bias the result.

A noticeable amount of literature focused on the acquisition of spoken word learning in adults. Gaskell and Dumay (2003) trained participants on spoken pseudowords which strongly overlapped with existing words (e.g. *'cathedruke'* derived from *'cathedral'*). The recognition task required participants to hear each novel word that presented along with its foil and indicated which of the two items was more familiar. Good explicit memory was shown after a single, concentrated exposure session. Lexicalization effects (the RTs to 'cathedral' was slowed by 46 ms) were absent immediately after exposure but arose after sleep occurs, without any further training. This suggests that new phonological information can be learned promptly, but full integration with existing knowledge requires a period of consolidation. This finding is consistent with learning new orthographic written words forms (Bowers et al., 2005) and developing picture naming connections by using a picture-word naming interference task that taps into orthographic and semantic processing (Clay, Bowers, Davis, & Hanley, 2007). Leach and Samuel (2007) propose a similar explanation with the result of recognition judgment and threshold discrimination tasks where *lexical configuration* (the set of factual knowledge associated with a word, e.g. the word's sound, spelling) can be developed with relatively few exposure to the word. Yet, *lexical engagement* (where a new word dynamically interacts with other lexicon representations) will require much more repetition of word exposure.

Using the same learning paradigm as Gaskell and Dumay (2003, 2007), Tamminen and Gaskell (2008) observed the lexicalization effect was clearly observable even 8 months after initial exposure. Although testing may act as a means of strengthening memory traces, periods without testing of up to 16 weeks did not eliminate competition effects. Thus, the competitive effects in these experiments cannot be explained as an episodic effect as the form of memory underlying these representations does not fade within a matter of days or weeks, as some episodic aspects of speech do (Goldinger, 1996).

Having seen a strong lexicalization effect in spoken word acquisition, Bowers et al. (2005) extended the result to understand the process of implicit written word learning. They introduced new words (BANARA) that were neighbours of familiar words that previously had no neighbours (BANANA). Repeated exposure to these new words made it more difficult to semantically categorize (natural or artefact) the familiar words. This shows evidence that competition between orthographically similar forms exerts an inhibitory effect on visual word identification. As mentioned earlier, in Salasoo et al (1985) study, participants had to read aloud the stimuli that they saw after the mask had appeared. In Bowers et al. (2005), the orthographic form of the new novel word does not provide any direct link to how the participants semantically categorize the original/base stimuli (e.g. the new orthographic pattern BANARA does not provide any information about how to classify BANANA in a semantic task). Accordingly, any impact of the new neighbours on classifying the targets would likely reflect lexical competition rather than some form of episodic influence. This resolved the plausible criticism of Salasoo, Shiffrin and Feustel study that learning in the threshold task might be episodic rather than lexical.

Acknowledging there is rapid learning in orthographic and spoken word learning, Chalmers and Burt (2008) took a further step to understand the role of phonological encoding skills in orthographic learning. In the training phase of the study, the orthography of each nonword was presented in the centre of the screen, with (P+) or without its pronunciation (P-). If present, pronunciation began at display onset. Participants were instructed to count the number of consonant clusters in the nonword (to encourage the processing of orthography) and to record their response by key press (*m* for more than 1, *n* for not more than 1). For the variation of semantic information, either the definition (S+) or the neutral phrase (S-) was presented with each nonword and participants were instructed to read the information silently. Learning was measured by an orthographic choice task. On each trial, a trained nonword (i.e. correct spelling) and a phonologically correct and orthographically acceptable distractor (i.e., incorrect spelling) were presented side by side. Participants were asked to judge which one was correct. The results showed that the provision of either phonological or semantic information during training improved spelling recognition. A similar result was obtained and extended by

Nelson, Balass, and Perfetti (2005) study. When the trained items were presented along with foils (half phonologically and half orthographically), they found that rare words that were trained with orthography and semantic meaning were learned better compared to words that were trained with phonology and semantic meaning. Taylor, Plunkett, and Nation (2011) also showed that pre-exposure to either phonology or semantics boosted the early stages of orthographic learning in artificial characters in the old-new decision task which trained artificial characters were mixed with untrained artificial characters.

Recognizing that the role of phonology is salient in word learning, McKague, Davis, Pratt, and Johnston (2008) manipulated the consonant/vowel structure of masked form primes to explore which element is more prominent in phonological learning. The method of mask priming was used to investigate word learning in British adults. In this procedure, a prime is presented briefly before the presentation of a target word. The results showed that items in the oral instantiation training preceded by the consonant-preserving form prime were recognized significantly faster than those preceded by the vowel-preserving form prime. Consistent with the consonant-frame hypothesis, orally instantiated novel words received significantly more facilitation from consonant-preserving form primes than from vowel-preserving forms.

Rastle, McCormick, Bayliss, and Davis (2011) took a different approach and examined the influence of orthography on spoken word production. They asked their participants to learn associations between spoken novel words and novel pictures. The following day, their participants learned the spellings of the novel words. Spelling-to-sound relationships were varied, with the spelling of the initial phoneme conforming to either regular English spelling-to-sound correspondences (e.g., the phoneme /k/ spelled *k*) or irregular ones (e.g., /k/ spelled *ch*). On the third day, participants had to name the pictures. Results showed that the novel words whose spellings were regular were named faster than those with irregular spellings, suggesting an influence of orthographic knowledge in spoken-word production. A similar result was obtained by Bürki et al. (2012) in novel French word learning.

Using a statistical learning paradigm, Breitenstein and Knecht (2002) tracked the progress of word learning as a function of time and exposure. Nonwords were assigned meanings by repeated pairing with a picture. The experiment also included a smaller proportion of incorrect nonword-picture pairs, thus requiring the participants to learn the correct pairings mainly by their statistical co-occurrence. Learning was measured by asking the participant whether each pair was a correct combination or not. Performance increased from chance level to 90% correct after 5 days of training and remained good 1 month after training. Another study (Breitenstein, Kamping, Jansen, Schomacher, & Knecht, 2004) replicated this finding and showed good performance even two months after the 5 training sessions.

1.7 Tapping into the process of orthographic learning

Though the aforementioned study were very helpful, they do not capture the rapid and automatic aspects of processing thought to be characteristic of skilled orthographic reading (Castles & Nation, 2008). The orthographic choice task involves presenting the reader with two alternative words with the same phonology at the same time, which is potentially confusing and which may actually disrupt the normal process of word recognition. While a spelling task does require access to complete specified representations, there has been debate within the field as to the degree to which this access process, and the associated representations, can be assumed to be the same as for those for visual word recognition (Holmes & Babauta, 2005). As some representation may be sufficient for recognition but be insufficient for reproduction of the word-specific knowledge required for accurate spelling. This meant a promising alternative to these standard tasks is required.

This thesis is concerned with the processes by which adults add new written words to their lexicons. It develops particularly on previous work by Weekes (1997) and Maloney, Risko, O'Malley, and Besner (2009). Weekes (1997) analysed the effect of word length on the speed with which adult readers of English can read aloud high frequency words (e.g., *car*, *film*, *spring*), low frequency words (e.g., *crab*, *freeze*, *sweep*) and invented nonwords (e.g., *colm*, *frip*, *slort*). Words and nonwords differing in length from 3 to 6 letters were interleaved and presented to participants in a random order.

Familiar words were read aloud more quickly than unfamiliar nonwords and while letter length had a strong effect on nonword naming speeds, the effect of length was smaller for low frequency words and not significant for high frequency words. Figure 1.3 *shows the main result of Weekes's (1997) study.*

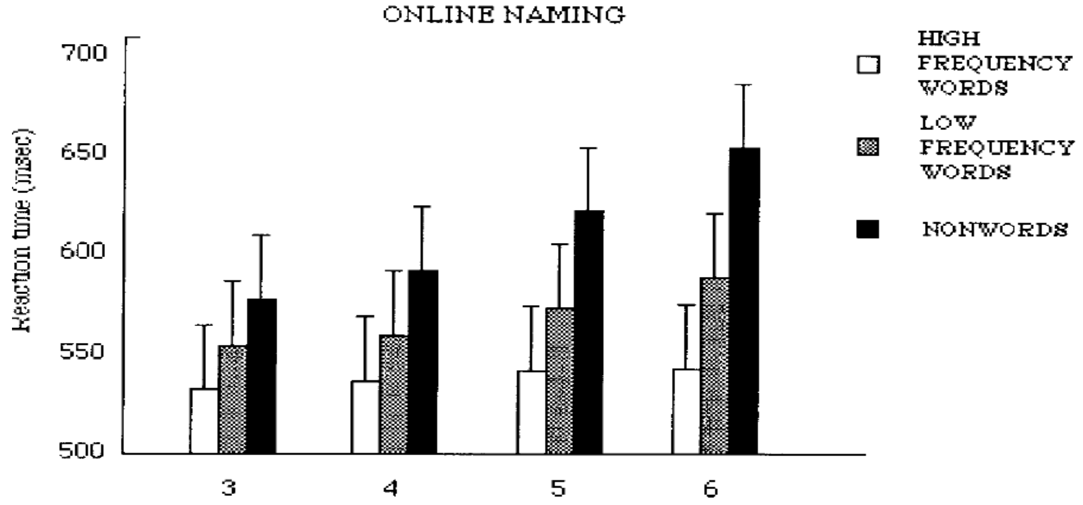

***Figure 1.3.* Naming RTs of high-frequency words, low-frequency words and nonwords (taken from Weekes (1997) Figure 1).**

Faster naming of words than nonwords has been documented in a number of studies that have not probed the interaction between lexicality and length in both lexical decision and reading aloud tasks in English (Johnston, McKague, & Pratt, 2004; Lupker, Brown, & Colombo, 1997; Rastle, Kinoshita, Lupker, & Coltheart, 2003; Scarborough, Cortese, & Scarborough, 1977), Italian (Pagliuca, Arduino, Barca, & Burani, 2008), and German (Ziegler et al., 2001). Differential effects of length on word and nonword naming resulting in a bigger lexicality difference for longer than shorter items has been reported in reading aloud in French (Juphard, Carbonnel, & Valdois, 2004; Valdois et al., 2006), English (Mason, 1978; McCann & Besner, 1987; Rastle & Coltheart, 1998), and German (Ziegler et al., 2001). Richards and Heller (1976) had obtained a similar interaction between length and lexicality using the exposure time required for successful identification of briefly-presented words and nonwords ("recognition thresholds") as their measure of performance rather than naming latencies. The larger effects of length on reading latencies for low than high frequency English words that Weekes (1997) noted have also been observed in a range of tasks including word naming and lexical decision tasks in several languages, including English (Cosky, 1976; Forster & Chambers, 1973;

Jared, Mcrae, & Seidenberg, 1990; Lee, 1999; Yap & Balota, 2009). An effect of length on naming latencies for lower frequency words may explain the consistent reports of significant, independent contributions of letter length to predicting RTs in large-scale analyses of adult word naming (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Cortese & Khanna, 2007; Cortese & Schock, 2013; Yap & Balota, 2009) for monosyllabic, disyllabic and multisyllabic words in English.

Taken together, these results imply that as novel, unfamiliar words become familiar through repeated exposure, naming latencies decrease and RTs to longer and shorter words converge. If that is true, it should be possible to simulate these dual aspects of visual word learning by using repeated exposure to familiarise participants with a set of initially-unfamiliar nonwords that vary in length. The result should be a progressive reduction in naming RTs and a convergence of RTs to shorter and longer items. That prediction was tested by Maloney et al. (2009) who presented Weekes's (1997) nonword stimuli (with a few minor modifications) to adult participants four times across four blocks of trials. Figure 1.4 shows the results. The effect of length was significant across the four presentations, reflecting faster overall responses to shorter than longer items. The effect of blocks was also significant, reflecting a speeding up of RTs with repetition. A significant length x blocks interaction in the by-participants analysis supported the indication in Figure 1.4 that the effect of length diminished across blocks as RTs to shorter and longer nonwords converged. This demonstrates different mechanisms are involved as words become more familiar and this relates to the account of modelling orthographic development--- a topic to which I now turn.

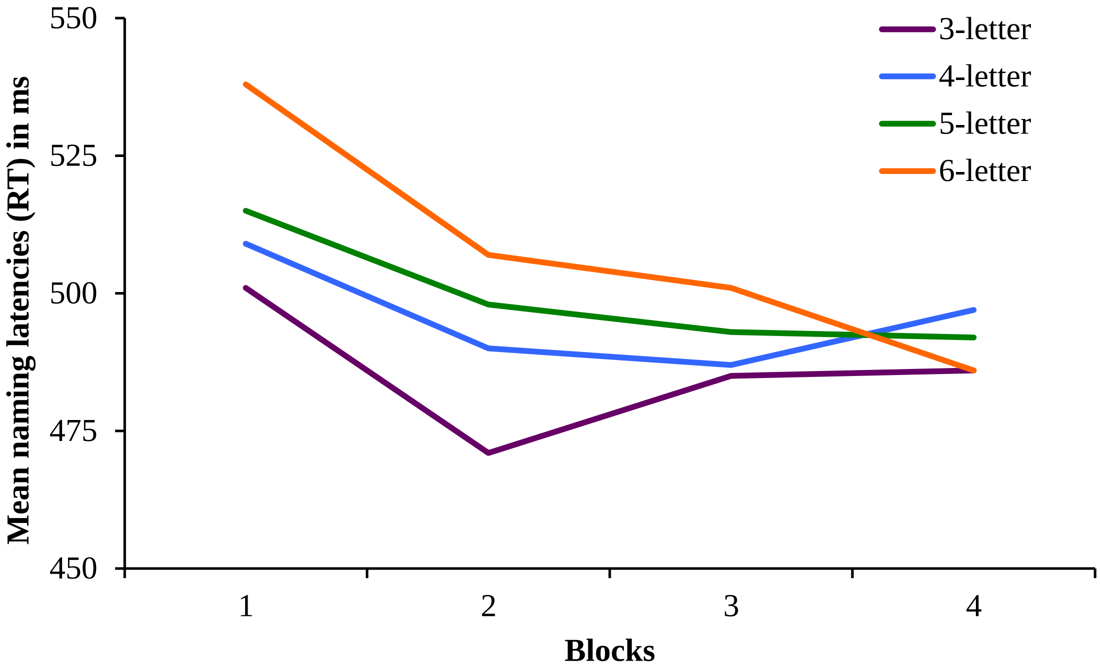

***Figure 1.4*. The mean naming RTs across four repetitions (blocks) in Experiment 1 of Maloney et al's (2009) study.**

1.8 How can the current computational models explain the mechanisms of word learning?

In the last two decades a number of successful computational models have been implemented to help understand the multiple processing levels of reading (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Grainger & Jacobs, 1996; Plaut, McClelland, Seidenberg, & Patterson, 1996). As this thesis mainly focuses on the development of mappings between orthography and phonology, that is, reading aloud, rather than recognizing and comprehending the meaning of words, the Dual Route Cascaded (DRC) model (Coltheart et al., 2001) will be treated as the main framework to explain orthographic learning in the following chapters. Other models, including PDP connectionist models (e.g. Plaut et al., 1996; Seidenberg & McClelland, 1989; Ševa, Monaghan, & Arciuli, 2009) and the Connectionist Dual Process model (CDP+) (Perry, Ziegler, & Zorzi, 2007, 2010; Zorzi, 2010) will be considered in Chapter 7 (the General Discussion chapter).

Based on Treisman (1961) suggestion that the mental lexicon stores representations of all known words, including their spellings, pronunciations and their meanings,

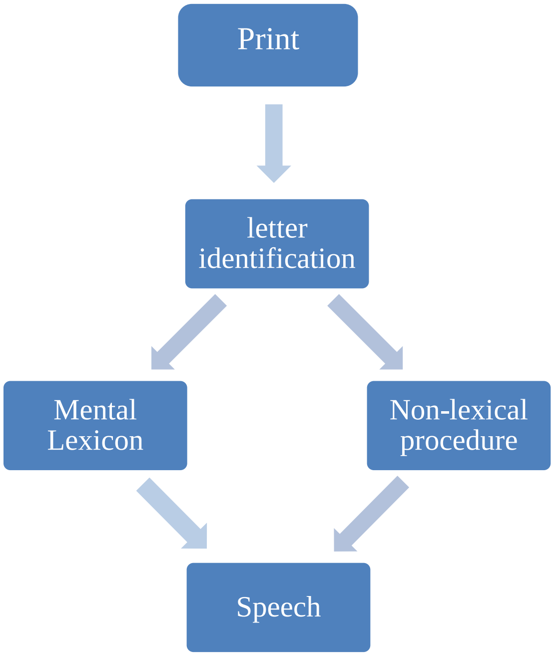

***Figure 1.5*. The basic dual route theory of reading aloud (modified from Coltheart et al. 2001)**

Figure 1.5 includes all these types of information in a single system. However, this version has been proven to be wrong. The result of neuropsychological research with people whose language has been disturbed by brain damage (Blazely, Coltheart, & Casey, 2005), compels researchers to adopt the view that these three forms of information about words are stored in three separate systems, as shown in Figure 1.6. Blazely et al. (2005) showed that in some people with dementia, knowledge of word meanings is severely impaired, but they can still perform the visual lexical decision task with normal accuracy (thus the orthographic input lexicon is intact) and can still read aloud irregular words with normal accuracy (thus the phonological output lexicon is intact as well). This show that only the semantic system is impaired is these patients.

There are two main assumptions of the updated DRC model in Figure 1.6. First, processing within the model is cascaded. This implies that as soon as there is activation in an early module, it flows to the next module instantly. Second, there are three transit routes in the model: the lexical semantic route, the lexical nonsemantic route, and the nonlexical grapheme-phoneme conversion (GPC) route. The general architecture of the DRC model can be seen in Figure 1.6. The model was named as *Dual Route* because the semantic system had not yet been implemented. The computation architecture of the DRC model is shown in Figure 1.7.

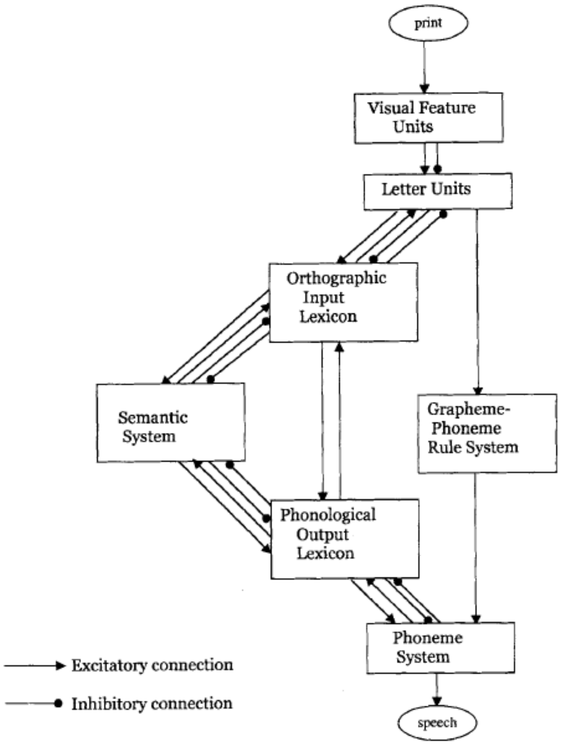

Figure 1.6. The computational architecture of the DRC model (taken from Coltheart et al., 2001)

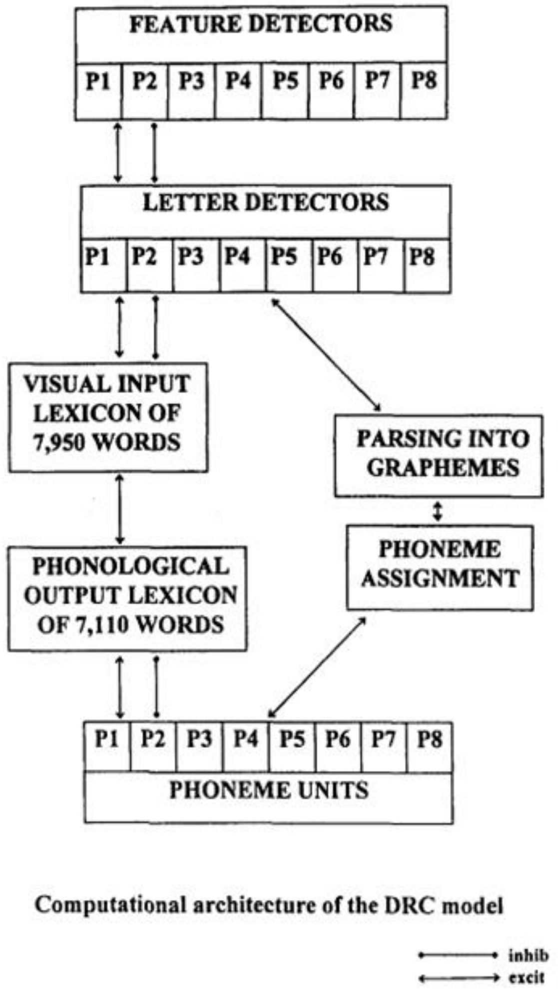

***Figure 1.7.* The computational architecture of the DRC model (taken from Dodd, Campbell and Worrall, 1996)**

1.8.1 The visual feature and letter units

The vocabulary that is stored in the orthographic lexicon of the DRC model contains 7,991 monosyllabic words based on the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993). Among those stored words, the longest words contain eight letters. Thus, the model has eight sets of letter detectors (one set for each position in the input string), and eight corresponding sets of feature detectors. For each of the letter detectors, there are 16 feature-present units and 16 feature-absent units based on the 16-stroke font used by Rumelhart and Siple (1974). When a set of letter strings is displayed to the model to be read aloud, it will first be presented at the feature level. All the features embedded by the first letter in the input string turned on their feature units, so do second letter and second set of feature units, and so on. This is named as the Cycle 0 in the process of the

DRC model that it sets all the units for visual features that are presented in the input string as 1 and sets all others to zero. Since units in the feature level are processed in parallel, the feature units in the letter string are activated simultaneously.

All the feature units in the first set of the letter strings are connected to all the letter units in the first set. Each position of a letter unit contains 27 units, one for every letter in the alphabet and one for the absence of any letter in that position in the input string. The feature units that are contained in the letter excite that unit while those that are not inhibit it. For instance, the feature ‘Horizontal in the middle’ excites letter units such as A, B, E, F, H, R and inhibits other letter units such as C and O. This is named as Cycle 1 of the process in the DRC model.

1.8.2 The non-lexical route

The non-lexical route consists of four components: the feature detection level, the letter units level, the grapheme-phoneme rule system and the phoneme system. The grapheme-phoneme rule system produces the pronunciation of letter strings (either a low-frequency word or nonword) through obeying the sub-lexical spelling-sound rules. Different from the lexical route, letters of the grapheme-phoneme rule system activates in a serial, left to right fashion. Activation of the second letter will not start until the processing of the first letter was complete. For instance, given a nonword ‘yacht’, the corresponding activation would be: Y -> /j/, A-> /æ/,C -> /s/, H ->/h/, T->/t/. Coltheart et al. (2001) suggest since the GPC route processes nonword in a serial order, the nonword letter length effect is an inevitable consequence of the process. In other words, since GPC translates letters serially, the time required to progress a nonword increases as the length of the nonword increases (see section 1.8.4 below).

1.8.3 The lexical route

The lexical route delivers the pronunciation of words based on word-specific knowledge. Other than the feature detection level and the letter unit level, this route contains three components: the orthographic input lexicon, the phonological output lexicon and semantic system. This is illustrated in the left side of Figure 1.7. The same kind of connection exists between the letter level and the orthographic input lexicon.

Thus, the unit for the letter A in the first set of letter units can excite the connection to every unit in the orthographic input lexicon containing a word starting with A, and inhibits all other units. In other words, the unit for the letter string APPLE in the orthographic input lexicon would excite the letter unit for A in the first set of letter units, and inhibit all other units in that set. This is referred to as Cycle 2 of the process in the DRC model.

Every unit in the orthographic input lexicon activates its representation in the phonological output lexicon directly. There are both excitatory and inhibitory connections between the phonological output lexicon to the phoneme level. As the longest eight letter monosyllabic word in the orthographic input lexicon contains only seven phonemes (certain letter represents two sounds), there are seven sets of representations in the phoneme level. The unit for the word 'APPLE' in the phonological output lexicon would excite the phoneme unit /a/ in the first set of phoneme representations, and inhibit all other phoneme representations in that set, then it would move on to the second and the third representation sets. This is equivalent to Cycle 3, 4 and 5 of the process in the DRC model. As the processing cycles progress, inhibitory and excitatory influences continue to flow upwards and downwards between layers until the reading-aloud response is ready. The inhibitory connections between the orthographic input lexicon and phonological output lexicon help to speed up the process of reading aloud. By the end of Cycle 4, some phoneme units will be activated, but extremely weakly. As processing continues, activation of some of the phoneme units will slowly rise. In the majority of circumstances, some of the phoneme units activated early in processing will be incorrect ones. Over time as phoneme activations continue to rise it is the correct phonemes that are the most activated. A reading response is considered to be ready when phonemes have reached a critical level of activation (set to .43 when the model is being used for simulating human reading aloud). The pronunciation generated by the model is taken to consist of the most highly activated phoneme within each of the eight sets of phoneme units (one set per position) that comprise the phoneme system.

The semantic system represents the meaning of a word while the lexicons compute the orthographic and phonological forms of the word. The lexical route can generate the pronunciation of all the words that are known by the computational model. Without the help of the lexical route, the computation model will not be able to pronounce an exception word which does not obey spelling-sound correspondence.

1.8.4 The transition from serial to parallel processing

Weekes (1997) and Maloney et al. (2009) explained their findings within the framework of the DRC model of visual word recognition (Coltheart et al., 2001). In that model, when novel words (or experimental nonwords) are encountered for the first time, they are read aloud through the application of grapheme-phoneme (letter-sound) conversion (GPC) rules which embody the most commonly-occurring correspondences between letters and sounds in English. This is consistent with the nonword naming result of Weekes (1997) and Maloney et al. (2009) in block 1 that there was a substantial length effect when the participants encountered the novel words for the first time. As novel words become familiar through repeated exposure, representations of the written forms of those words are created within the orthographic input lexicon while representations of their spoken forms are created within the phonological output lexicon. This is in line with the result of Maloney et al. (2009) that the length effect of the novel words reduced from block 2 onwards.

The ability of the DRC model to simulate the interaction between lexicality and letter length was reported by Coltheart et al. (2001, p. 239; see also Perry & Ziegler, 2002; Perry, Zeigler, & Zorzi, 2007; 2010). Coltheart et al. (2001) further demonstrated the interaction between letter length effect and lexicality in human data can be simulated by the DRC model. Processing along the nonlexical route does not begin to operate until cycle 10. Without this time frame after the lexical route begins to operate, the model would have serious difficulty in reading aloud irregular words. When Cycle 10 is reached, the nonlexical route translates the first letter of the string into its phoneme using the appropriate grapheme-phoneme rule. Every 17 cycles, the GPC system moves on to consider the next letter of the nonword, translate it to a phoneme, and activate that phoneme in the phoneme system. Thus, with the letter string BRUP, the GPC system has no input until cycle 10, deals with just B until cycle 27, deals with just BR from cycle 28 to cycle 44, then BRU until cycle 60, BRUP until cycle 76 and so on.

Figure *1.8* shows the result of length effect and lexicality on naming latencies from human readers and DRC model. There was clearly a significant effect of length for nonwords but not for words from the human data. Similarly for the DRC model, ANCOVA result (neighborhood size as covariate) showed there was only an effect of length for nonwords but not for words. As mentioned earlier, since the GPC route has to process the nonword letter string serially, while the lexical route processes words in parallel, the length effect can only be observed for the naming latency of nonwords.

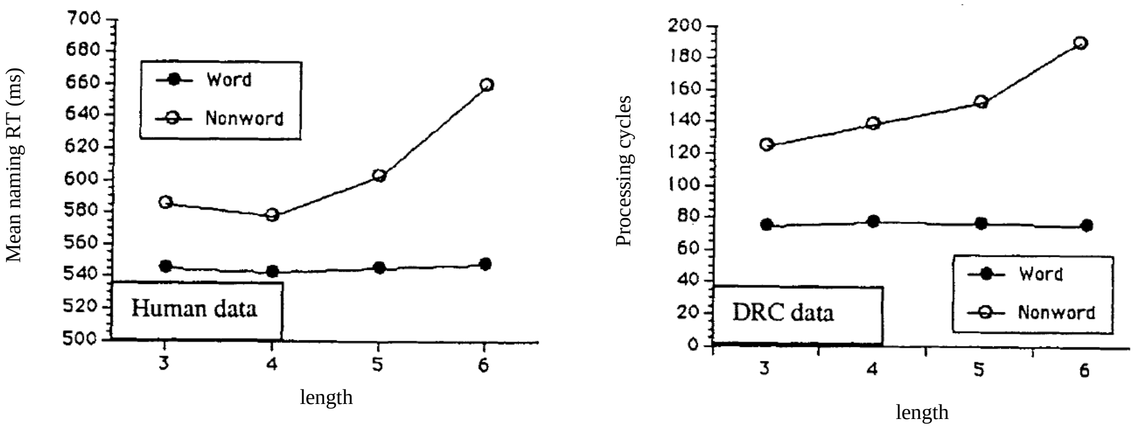

***Figure 1.8.* Effects of length and lexicality on naming latencies from human readers and DRC model (taken from Coltheart et al., 2001, Figure 10).**

Not only can the DRC model stimulate the lexicality effect in human data for reading aloud tasks, it also replicates the frequency effect in human data for lexical decision task. The DRC model was built to provide a YES response if 1) any entry in the orthographic lexicon has been reached to a certain amount (0.69), 2) if the sum of the activations of all the entries in the orthographic lexicon has reached 10 which met the criterion of the 'fast-guess' mechanism. The DRC model was built to provide a NO response if the processing cycles had elapsed and a YES decision has not yet been made. Based on the human data result from Andrews (1989, 1992), Coltheart (2001) found a significant effect of word frequency in ANOVA analysis. There is also a significant interaction of frequency and neighborhood size, but only for low-frequency words, which replicated the human data from Andrews (1989, 1992). The exact result was obtained from the YES latencies of the DRC model by the number of cycles it took the DRC model to provide an answer. Figure 1.9 shows the mean Yes latencies from the human data (Andrews 1989, 1992) and the DRC model's mean correct Yes. The process of how the frequency and neighborhood size affect word learning will be further explored in Chapter 4 in which high- , low-frequency words and nonwords are included in the experiment.

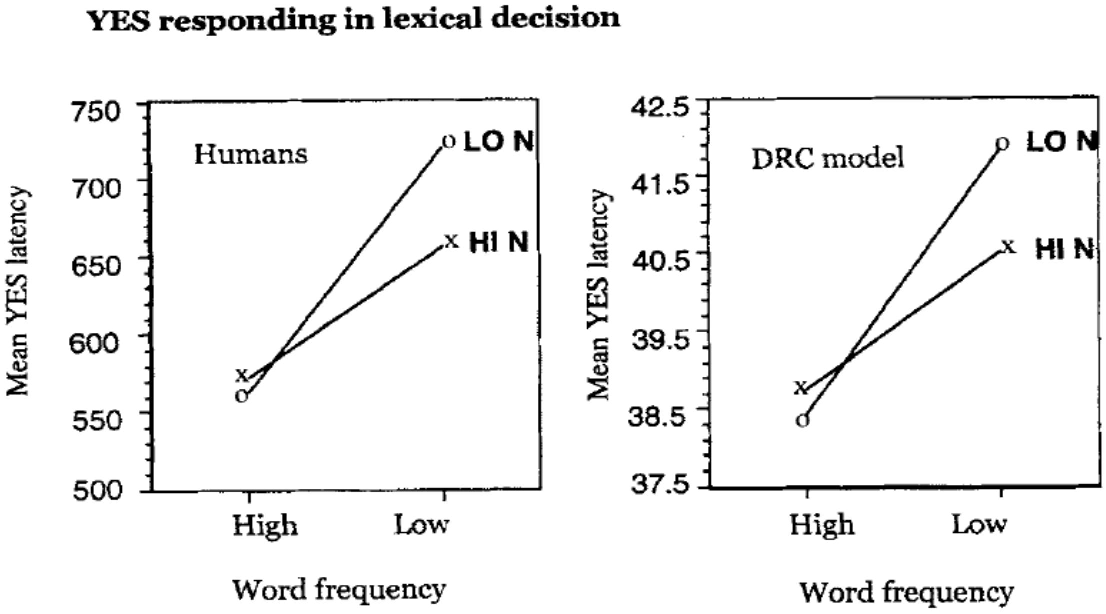

Figure 1.9. **Effects of frequency on reaction latencies from human readers and DRC model (adopted from Coltheart et al., 2001, Figure 8).**

One can also link the mechanism of the dual route model to the predictors of word learning that is mentioned in section 1.4. Based on section 1.4.1, if a child has good alphabetic and decoding skills, then she/he will have a normal-for-age development in the visual feature units, letter units and good processors of the non-lexical route. If a child has good implicit phonological processing skills, this will speed up the progression in all levels of the lexical and non-lexical route. If a child has good vocabulary skills, this means that she/he will have a strong and comprehensive orthographic input lexicon for her/him to relate to while they are learning new words. Finally, if a child has a solid working memory span, this will help to retain the information that they acquire in the orthographic input lexicon and phonological output lexicon.

1.9 Conclusions and thesis outline

While there is a rich and solid foundation on the role of orthography and phonology in word learning, further studies are needed to understand how new words are learned implicitly as readers' mental lexicons grow (Castle & Nation, 2008). Using a novel learning paradigm, the present thesis therefore brings together several complementary approaches to understand the process of orthographic word learning.

To address the research questions, the present thesis contains five exploratory investigations across five chapters. Chapter 2 demonstrates the word learning paradigm that will be used throughout this thesis and explores how many exposures are required before a new word can build lexical entries in the mental lexicon and be processed in a unitized way. The memory retention of these learned representations is also investigated in Chapter 2. Chapter 3 seeks to understand the role of phonology in orthographic word learning. Chapter 4 then aims to investigate how the newly learned items integrate with existing knowledge in the mental lexicon with high- and low-frequency words. Finally, Chapters 5 and 6 try to bring these theories in a more applied context to understand orthographic word learning in adults with dyslexia and in bilingual speakers.

2 Chapter 2: Visual word learning in skilled readers of English

2.1 Introduction

This chapter is concerned with the processes by which adults add new written words to their lexicons. It builds in particular on previous work by Weekes (1997) and Maloney et al. (2009). The word learning paradigm that will be used throughout this thesis will be addressed. The number of exposures that are required before a new word can built lexical entries in the mental lexicon and its retention will also be discussed.

Word recognition grows in such remarkable speed that, by the end of eighth grade, children who learn to read English know and recognise over 80,000 words (Adams, 1990). Beginning readers must develop a system of mappings or correspondences between the letters or graphemes of written words and the phonemes of spoken words Ehri (1992), and it is established that this alphabetic decoding system is supported by phonological skills (Bradley & Bryant, 1983; Hulme & Snowling, 2014). As children grow up, the process of word learning never stops: Nation and Waring (1997) estimate that the receptive vocabulary size of a university-educated native English speaker is around 20,000 base words.

As mentioned in Chapter 1, Weekes (1997) found that familiar words were read aloud more quickly than unfamiliar nonwords and while letter length had a strong effect on nonword naming speeds, the effect of length was smaller for low frequency words and not significant for high frequency words. Differential effects of length on word and nonword naming resulting in a bigger lexicality difference for longer than shorter items has also been observed (e.g. Juphard et al., 2004). Maloney et al. (2009) also found that the effect of length on nonword reading diminished across blocks as RTs to shorter and longer nonwords converged with repetition. The result of these studies imply that as novel, unfamiliar words become familiar through repeated exposure, naming latencies decrease and RTs to longer and shorter words converged. As mentioned in section 1.8.4 in Chapter 1, this process from serial to parallel processing in English novel words can be explained by the DRC model (Coltheart et al., 2001) as English novel words changed from processing in the majority by the non-lexical route to the lexical route. Thus, by block 4 of the training session of Maloney et al. (2009), though the non-lexical route cannot stop its contribution towards the novel words naming tasks, given that the lexical route operates very quickly, verbal response is delivered by the lexical route before the non-lexical route is able to produce any responses. On the basis of their findings, Maloney et al. (2009) suggested that skilled readers can create entries for new words in the orthographic input lexicon and phonological output lexicon after just 3 or 4 presentations.

Similar estimates of the number of presentations required to created lexical representations have come from other studies that have employed a variety of methods to analyse word learning in both adults and children. Using a threshold recognition task, Solomon and Postman (1952) asked participants to recognize novel words that were ‘buried’ (masked) by reading the stimuli aloud. They found an effect of previous presentations on the exposure time required for adults to identify 7-letter nonwords correctly. Duration thresholds fell rapidly from first to third presentation then reduced more slowly thereafter. Salasoo et al. (1985) found lower recognition thresholds for words than nonwords when the stimuli were presented to adults for the first time. Thresholds then reduced with repetition for both words and nonwords, but more so for nonwords than words. Thresholds asymptoted after around five presentations after which the difference in thresholds between words and nonwords was no longer detectable.

Studies of word leaning in normally-developing children have suggested similar estimates of the number of exposures to a novel word required to create orthographic and phonological representations. In Hogaboam and Perfetti (1978) study, children in third grade of schooling (8-9 years of age) repeated nonwords spoken by the experimenter with or without the spelling of the nonword presented for the child to look at. Each nonword was presented 3, 6, 12 or 18 times over three sessions on three consecutive days. On the fourth day the children were asked to read aloud all the nonwords presented in written form in addition to a set of untrained items. Naming latencies were quicker to trained than untrained nonwords, and shorter following training with the orthography of the nonword presented than after purely auditory training. For more skilled readers, the benefits of training exposures were as great following three exposures as following 6, 12 or 18 exposures.

Reitsma (1983) trained Dutch children aged 7-8 years to read versions of familiar words that had been re-spelled in a way that preserved the word's pronunciation but changed its presence (i.e., "pseudohomophones" equivalent to re-spelling the English word *keep* as *keap*). The re-spelled versions of the words were presented either four or eight times during training. Three days later the children were asked to read aloud correctly-spelled versions of the trained words along with untrained, control words. The words that had been trained by reading aloud versions that preserved the phonology but changed the orthography were read aloud faster than that untrained control words. The benefits of prior training were as strong following four presentations in training as following eight. Similar indications that between 3 and 5 presentations of novel words are sufficient to create new, functioning representations in children can also be found in the studies by Ehri and Saltmarsh (1995), Manis (1985) and Share (1999).

This chapter reports three experiments investigating visual word learning in skilled, adult readers of English. Experiment 1 represents a replication and extension of Maloney et al. (2009). Participants read aloud 12 4-letter and 12 7-letter nonwords that were interleaved and displayed in different random orders across 10 blocks. The instructions were to read each nonwords as quickly and as accurately as possible when it appeared on the computer screen. On the basis of Weekes (1997), Maloney et al. (2009) and other studies it is expected to see a substantial effect of letter length on naming latencies the first time the nonwords were presented (block 1). It is hypothesized that RTs would reduce across blocks and that RTs to shorter and longer items would converge over 3 to 5 presentations as lexical representations are created and reading switched from nonlexical to primarily lexical. Experiment 2 then investigated the extent to which any reduction of RTs with repeated exposure and convergence of RTs to shorter and longer items is a consequence of item-specific training or more general improvement on the task while Experiment 3 examined whether the effects of 10 presentations of nonwords in one session would be detectable in performance on the same nonwords a week later.

2.2 Experiment 1: learning through repeated exposure

2.2.1 Method

2.2.1.1 Participants

Participants were 25 undergraduate students of the University of York (12 male, 13 female) with a mean age of 20.16 years (S.D. = 2.01; range 18 - 28). All were native speakers of English with normal or corrected-to-normal vision and no history of reading or language problems. Participants received either course credit or a small payment. This and the other experiments reported here were approved by the Ethics Committee of the Department of Psychology, University of York.

2.2.1.2 Materials

90 monosyllabic four-letter nonwords and 89 bisyllabic seven-letter nonwords were generated based on the WordGen nonword generation program (Duyck, Desmet, Verbeke, & Brysbaert, 2004) on the basis of the CELEX lemma database (Baayen et al., 1993; Baayen, Piepenbrock, & Van Rijn, 1995) and the Lexique database (New, Pallier, Brysbaert, & Ferrand, 2004). To generate a nonword, the program randomly arranges a string of selected letters and verifies whether the letter string is an existing word in the two lexical database of the particular language. Then, every constraint is processed (including, the length of nonwords and the bigram frequency range), and as soon as one of them is violated the random letter string is rejected and the operation starts all over again until a letter string fits all the constraints.

The sets of four-letter and seven-letter nonwords were matched on initial letters, and bigram frequency. None of the nonwords has a written or spoken form that is similar to a real word (i.e. the sets contained no ‘pseudohomophones’). To reduce problems with voice key activation, none of the nonwords began with a voiceless fricative (‘f’, ‘s’, ‘sh’, or ‘th’). Twenty one participants took part in a pilot study in which they were asked to pronounce the 179 nonwords one at a time as they were shown on screen. RTs shorter than 200 ms or longer than mean plus 2.5 SDs for each participant across four- and seven- letter items were regarded as outliers and removed from the analyses of RTs. The results of 19 participants which had accuracy above 75 percent were taken into further analysis. Sixteen items that had accuracy below 14/19 (73 percent) were deleted from the list. Based on the result of the pilot testing, one set of nonwords (24 nonwords, Set A) which were matched on accuracy (all above 90 percent) and on initial letters (12 different letters to make the nonwords as different as possible) were chosen to be the experimental items. All the stimuli of Set A is shown in Appendix 6. Reading speed was matched separately for 4- and 7-letter nonwords. The range, mean and standard deviation of the four- and seven- letter experimental items were shown in Table 2.1. An addition of sixteen nonwords (8 four-letter, 8 seven-letter) were chosen for practice trials prior to the main experiment.

2.2.1.3 Procedure

After completing a consent form, participants were given practice on the task which involved reading 8 4-letter and 8 7-letter nonwords presented in a random order. The experimental task was then given. Participants were seated approximately 60 cm from a computer screen on which the nonwords were displayed in black, lower case letters on a white background. The nonwords were presented in 18-point Times New Roman font with a height on the screen of approximately 10 mm. Each trial consisted of a centrally-presented fixation cross displayed for 1,000 ms, followed by the nonword stimulus for 2,000 ms then a blank screen for 1,000 ms before the next trial began. Participants were instructed to read each nonword aloud as quickly and as accurately as possible. The 24 nonwords were presented once in a random order. Participants were informed when the block was complete and pressed the space bar on a computer keyboard to initiate the next block when they were ready to continue. This process was repeated across 10 blocks with the stimuli being presented in a different random order in each block. Participants wore headphones with a high-sensitivity microphone connected to a voice key that was linked to the computer. Presentation of the stimuli and recording of naming latencies was controlled by E-prime experiment generator software (Schneider, Eschman, & Zuccolotto, 2002). No feedback was provided but the experimenter noted any trials in which the participant misread a nonword, hesitated or made a false start or other form of error.

Table 2.1. ***Mean and standard deviation of bigram frequency, neighborhood size, reading speed and accuracy of the four and seven-letter nonwords of Set A from the pilot study.***

	Nonwords	
	4-letter	7-letter
Bigram frequency		
Mean	1910	2327
S.D.	1391	834
Log Bigram frequency		
Mean	3.14	3.34
S.D.	0.40	0.16
Neighbourhood size		
Mean	6.75	0.17
S.D.	3.62	0.39
Phonemes		
Mean	3.67	5.83
S.D.	0.49	0.83
Reading speed (in ms) in pilot study		
Mean	631	725
S.D.	43	40
Range	573 - 733	669 - 782
Naming accuracy in pilot study		
Mean (max = 19)	17.5	17.5
S.D.	0.67	1.51
Range	16 - 18	14 - 18

Note. S.D. = standard deviation

2.2.2 Result

Only RTs for correct responses were analysed. Naming errors, hesitations and failures to activate the voice key accounted for 3 trials (0.05% of the total). RTs shorter than 200 ms or longer than mean plus 2.5 SDs in each block for each length group were regarded as outliers and removed from the analyses of accuracy and RTs. This led to the loss of a further 74 RTs (1.2% of the total), leaving 5923 RTs (98.7% of the total) for analysis. The mean RTs (with standard deviation) in each block for four-and seven-letter nonwords are shown in Table 2.2 along with the final accuracy (maximum = 12) in each condition.

Table 2.2. ***Mean latencies of correct, trimmed responses, standard deviation (SD), and per cent correct responses for 4- and 7-letter nonwords in blocks 1 to 10 of day 1 and day 7 in Experiment 1.***

Blocks	1	2	3	4	5	6	7	8	9	10
4 letters										
Mean RT	588	513	505	498	483	482	483	484	474	474
S.D.	96.8	76.4	87.8	71.8	58.3	51.5	58.8	64.6	66.2	61.2
Mean Acc.	11.92	11.84	11.96	11.84	11.84	11.84	11.80	11.92	11.84	11.96
S.D.	0.28	0.37	0.20	0.37	0.37	0.37	0.41	0.28	0.37	0.20
% correct	99.3	98.7	99.7	98.7	98.7	98.7	98.3	99.3	98.7	99.7
7 letters										
Mean RT	668	573	552	515	512	503	490	495	487	488
S.D.	112	52	85.4	80	55.4	53.7	53.9	53.2	63.5	79.9
Mean Acc.	11.76	11.88	11.68	11.84	11.64	11.80	11.96	11.92	11.84	11.84
S.D.	0.44	0.33	0.48	0.37	0.57	0.41	0.20	0.28	0.37	0.37
% correct	98.0	99.0	97.3	98.7	97.0	98.3	99.7	99.3	98.7	98.7

Note. RT = Reaction time (naming latency) in ms; S.D. = standard deviation; Acc. = Accuracy

2.2.2.1 Accuracy

Accuracy was generally very high (overall mean 98.7% correct and never below 97.0% in any condition). Given the high accuracy levels, nonparametric tests were employed. Wilcoxon matched-pairs signed ranks tests found no significant differences between the overall accuracy of responses to 4- and 7-letter nonwords, $W(25) = 68.00$, $Z = -1.40$, $p = .162$, or between levels of accuracy in blocks 1 and 10, $W(25) = 20.00$, $Z = 1.13$, $p = .257$.

Figure *2.1* shows the pattern of accuracy for correct, trimmed responses across blocks.

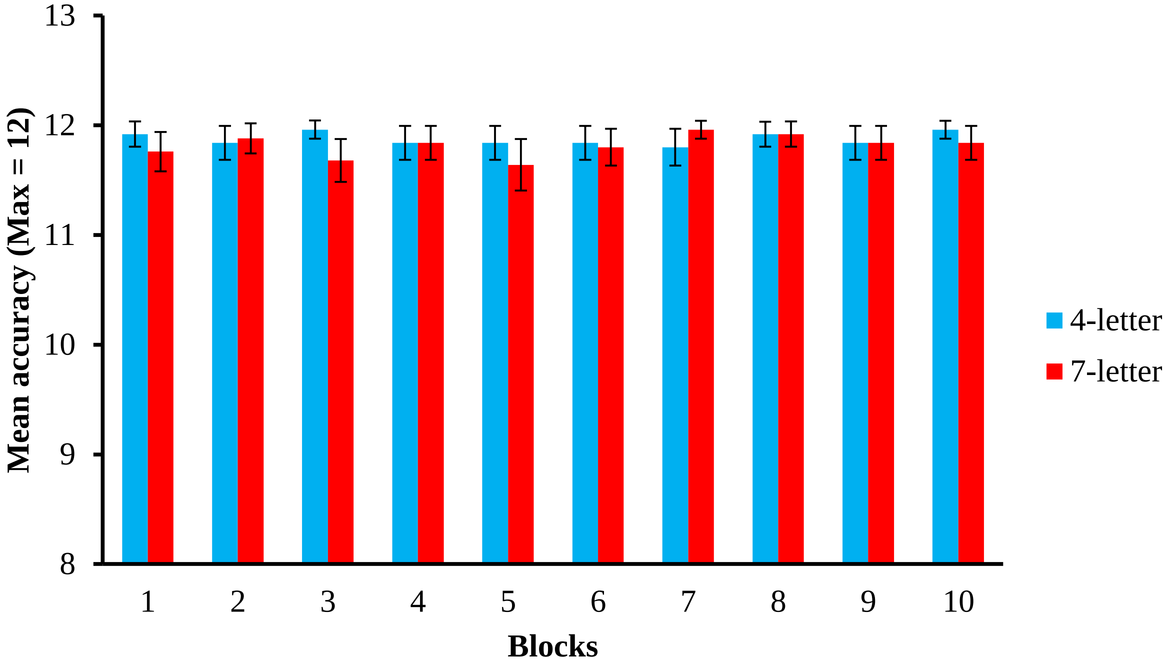

***Figure 2.1.* The accuracy of naming 4- and 7-letter nonwords in Blocks 1 to 10 in the trained and untrained conditions. Error bars show 95% confidence intervals.**

2.2.2.2 Naming latencies (RTs)

Figure 2.2 shows the pattern of RTs for correct, trimmed responses across blocks. The figure shows a reduction in naming latencies across the first 6 or 7 blocks after which RTs approach asymptotic levels. The general reduction in RTs is accompanied by a decline in the effect of length, with a large difference between 4- and 7-letter nonwords in block 1 reducing to a very small difference from around block 7 onwards.

RTs were analysed using a two-way ANOVA with Blocks (1-10) and Length (4 vs. 7 letters) as factors[1]. When Mauchly’s test of sphericity was significant, the Greenhouse-Geiger correction was applied. Bonferroni-corrected *t*-tests were used when

---

[1] Raaijmakers , Schrijnemakers, and Gremmen (1999, p. 426) argued that "when the materials have been matched on a number of variables or when the lists are counterbalanced over different groups of subjects ... the simple subject analysis will be correct". Accordingly, only the by participants (F_1) analysis will be presented and discussed.

pairwise comparisons were required. Full details of the statistical analyses can be found in the section 1.1.1 of Appendix 1 where effect sizes are reported in terms of the partial eta squared statistic (η_p^2). The main findings are summarized here.

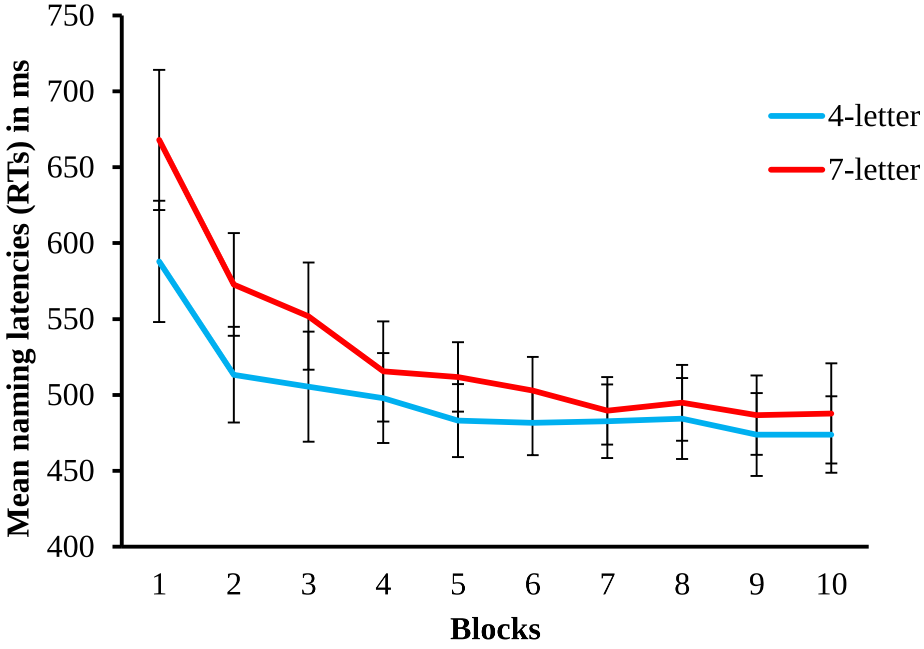

Figure 2.2. **The naming reaction times (RTs) for 4- and 7-letter nonwords in Blocks 1 to 10. Error bars show 95% confidence intervals.**

The main effect of Blocks was significant, with overall RTs becoming faster across blocks. The main effect of Length was also significant, with faster overall responses to 4- than 7-letter nonwords. A significant interaction between Length and Blocks reflected the fact that the difference between RTs to 4- and 7-letter items reduced across blocks from 110 ms in block 1 to 14 ms in block 10. In pairwise comparisons, the effect of length was significant in blocks 1, 2, 3, 5 and 6, but not in blocks 7 to 10.

2.2.3 Discussion

Pre-selection of the items for Experiment 1 on the basis of the pilot study meant that accuracy of reading the nonwords was high throughout. Ceiling effects meant that there was no detectable influence of length or blocks on accuracy; also that very few trials were lost from the RTs analysis.

Naming latencies to nonwords seen for the very first time in block 1 were 588 ms for 4-letter nonwords and 668 ms for 7-letter nonwords. That compares with 575 ms and 666 ms for the 4- and 6-letter nonwords in Weekes (1997). The means for the 4- and 6-letter nonwords in block 1 of Maloney et al.'s (2009) Experiment 1 were somewhat faster (509 ms and 538 ms respectively). The difference of 110 ms in mean RTs to 4- and 7-letter nonwords in the present experiment illustrates the well-established effect of length on naming speed for unfamiliar nonwords (cf. Juphard et al., 2004; Mason, 1978; Valdois et al., 2006; Weekes, 1997; Ziegler et al., 2001).

RTs became shorter across blocks as the nonwords became familiar. This was particularly true for the longer nonwords. Mean RTs for 4-letter nonwords reduced by 114 ms across the 10 blocks of training while the mean RTs for 7-letter nonwords reduced by 180 ms. The result was the convergence of RTs to shorter and longer nonwords that is very apparent in Figure 2.2. In fact, the effect of length became nonsignificant after block 6. The results for the first 4 blocks mirror the findings of Maloney et al. (2009), with RTs becoming faster and length effects diminishing across blocks.

In dual-route terms (Coltheart et al., 2001) the present results would be explained in terms of the nonwords being converted from orthography to phonology using the nonlexical route when they are shown for the first time in block 1. Over the course of the first few blocks, representations are created in the orthographic input lexicon and the phonological output lexicon which enable lexical reading to develop. The speeding up of naming responses and the convergence of RTs to shorter and longer nonwords reflect the change of processing mainly from nonlexical to lexical reading. From around block 7 onwards, lexical reading is established, the nonwords are read rapidly and the effect of length is no longer significant. This matches the indications in studies of both adult and child readers (e.g. Hogaboam and Perfetti's, 1978) that 4 or 5 presentations of novel words (nonwords) is sufficient to create representations that facilitate rapid identification and more parallel processing of component letters.

This account, like all the other account of visual word learning, assumes that the effects of repeated exposures are due to experience with the specific, repeated items (for example, repetition causes lexical entries to be formed for the novel items that facilitate subsequent recognition and naming of those items and only those items). Experiment 1 did not, however, include sets of nonwords that were tested at the beginning and end of training on the repeated set to see if any of the benefit of repeated naming generalizes to non-repeated items. The same is true of other studies that have examined the effects of repetition on responses to novel words. The way to assess that possibility is to compare RTs for items that are repeated across blocks with RTs to items that appear only before the start of training (block 1) or only at the end of training (block 10). That is accomplished in Experiment 2.

2.3 Experiment 2: item-specific or general learning?

Three sets of nonwords (B, C and D) were created, with each set containing 12 4-letter and 12 7-letter items (as in Experiment 1). The sets were matched on initial letters and phonemes, and on naming RTs from the pilot study. Each participant received one set of nonwords in all 10 blocks of the experiment. A second set was presented in block 1 only, randomly interleaved with the to-be-repeated set while a third set was presented in block 10 only, again randomly interleaved with the repeated nonwords. The three sets of nonwords were counterbalanced across conditions and participants so that each set presented equally often as repeated items or as non-repeated items in block 1 only or block 10 only. Assuming that performance on the repeated set would follow the same pattern as in Experiment 1, the question of interest was how RTs to the non-repeated (untrained) set in block 10 would compare with RTs to the equivalent set in block 1.

2.3.1 Method

2.3.1.1 Participants

Participants were 24 undergraduate students of the University of York (12 male, 12 female) with a mean age of 19.71 years (S.D. = 1.37; range = 18 - 23). All were native speakers of English with normal or corrected-to-normal vision and no history of reading or language problems. None had taken part in Experiment 1. Participants received either course credit or a small payment.

2.3.1.2 Materials

The experimental stimuli were three sets of nonwords (Sets B, C, and D), with each set containing 12 4-letter items and 12 7-letter items. The 12 4-letter and 12 7-letter nonwords in each set began with 12 different consonant letters. The range, mean and standard deviation of the 4- and 7- letter experimental items from the pilot study are shown in Table 2.3. All of the nonwords had accuracies above 90% in the pilot study. None began with a voiceless fricative. All of the stimuli of Sets B, C, and D are shown in Appendix 6.

2.3.1.3 Procedure

After completing a consent form, participants were given practice on the task which involved reading 8 4-letter and 8 7-letter nonwords presented in a random order. The experimental task was then given. Block 1 of the experiment contained nonwords from two sets, interleaved in a random order. One of the sets was then repeated in blocks 2 to 9, using a different random order in each block. In block 10, the set that had been presented throughout blocks 1 to 9 for that participant was presented again, but interleaved with a third set of nonwords in a random order. The result was that one set of items (B, C or D) was presented in all 10 blocks of the experiment, one set was presented in block 1 only, and one set was presented in block 10 only. The assignment of sets to conditions was counterbalanced across participants. Participants were instructed to read every nonword aloud as quickly and as accurately as possible when it appeared on the screen. Other details of the procedure were the same as for Experiment 1.

2.3.2 Result

Only RTs for correct responses were analysed. Naming errors, hesitations and failures to activate the voice key accounted for 42 trials (0.6% of the total). RTs shorter than 200 ms or longer than mean plus 2.5 SDs in each block for each length group were regarded as outliers and removed from the analyses of accuracy and RTs. This led to the loss of a further 67 RTs (1.0% of the total), leaving 6803 RTs (98.4% of the total) for analysis. The mean RTs (with standard deviation) in each block of the two conditions for four-and seven-letter nonwords are shown in Table 2.4 along with percent of correct trials in each condition.

Table 2.3. ***Mean and standard deviation of bigram frequency, neighborhood size and reading speed of the four and seven-letter nonwords of Set B, C and D from the pilot study.***

	Nonwords					
	Set B		Set C		Set D	
	4-letter	7-letter	4-letter	7-letter	4-letter	7-letter
Log bigram frequency						
Mean	3.19	3.35	3.09	3.29	3.19	3.34
S.D.	0.35	0.11	0.30	0.13	0.23	0.15
Neighborhood size						
Mean	6.08	0.08	6.58	0.08	5.67	0.17
S.D.	3.96	0.29	3.90	0.29	3.60	0.58
Phonemes						
Mean	3.67	6.00	3.58	5.83	3.75	5.92
S.D.	0.49	0.74	0.67	0.58	0.45	0.79
Reading speed (in ms) in the pilot study						
Mean	634	745	637	741	636	743
S.D.	38	77	47	53	55	71
Range	573 - 695	647 - 899	588 - 733	669 - 811	529 - 738	641 - 836
Naming accuracy in the pilot study						
Mean (max = 19)	17.75	17.50	17	17.08	17.42	17.42
S.D.	0.97	1.67	0.79	1.16	1.16	1.51
Range	17-19	14-19	16-18	16-19	15-19	15-19

Note. S.D. = standard deviation

Table 2.4. ***Mean latencies of correct, trimmed responses, standard deviation, and per cent correct responses for 4- and 7-letter trained nonwords in blocks 1 to 10 and for untrained nonwords in blocks 1 and 10 only in Experiment 2.***

Untr.						***Trained***						***Untr.***
Blocks	*1*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*10*
4 letters												
Mean RT	581	595	533	515	508	501	506	506	521	508	520	549
S.D.	58.1	78.3	62.6	64.3	72.1	75.1	61.1	61.9	69.8	75.2	60.1	78.7
Mean Acc.	11.8	11.8	11.9	11.8	11.9	11.8	11.7	11.9	11.7	11.8	11.8	11.8
S.D.	0.44	0.53	0.34	0.66	0.45	0.38	0.46	0.34	0.56	0.38	0.51	0.53
% corr	97.9	97.9	99	98.3	99	98.6	97.6	99.0	97.2	98.6	98.3	97.9
7 letters												
Mean RT	693	704	582	545	531	525	522	517	526	529	544	629
S.D.	127.2	134.1	90.3	82.7	71.4	87.0	68.0	70.0	71.6	81.5	77.7	15.3
Mean Acc.	11.8	11.9	11.8	11.8	11.8	11.9	11.8	11.8	11.9	11.8	11.8	11.8
S.D.	0.38	0.34	0.38	0.61	0.48	0.28	0.51	0.38	0.34	0.51	0.48	0.38
% corr	98.6	98.6	98.6	97.9	98.6	99.3	98.3	98.6	99.0	98.3	98.6	98.6

Note. RT = Reaction time (naming latency) in ms; SE = standard error; Untr = Untrained;
% corr = precent of correct

2.3.2.1 Accuracy

Accuracy was very high (overall mean 98.4% correct and never below 97.9%) in any condition. Ceiling effects meant that there were no significant differences between accuracy to 4- and 7-letter nonwords for the trained set across all 10 blocks, $W(24) = 93$, $Z = 0.79$, $p = .429$. There was also no significant difference in overall levels of accuracy to trained nonwords in blocks 1 and 10, $W(24) = 18.00$, $Z = 0.00$, $p = 1.00$, or to trained and untrained items in block 1, $W(24) = 34.00$, $Z = -0.42$, $p = .675$, or block 10, $W(24) = 29.50$, $Z = -0.33$, $p = .745$. Figure 2.3 shows the pattern of accuracy for correct, trimmed responses across blocks of the trained and untrained items.

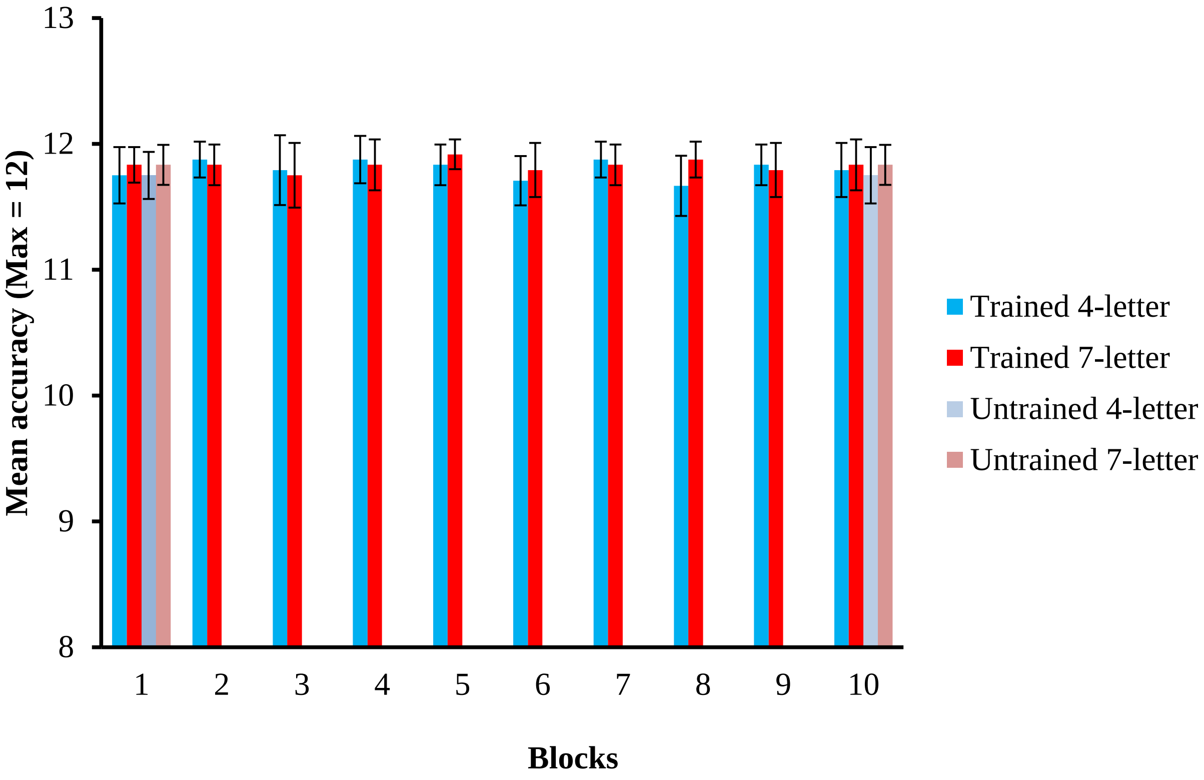

***Figure 2.3.* The accuracy of naming 4- and 7-letter nonwords in Blocks 1 to 10 in the trained and untrained conditions.**

2.3.2.2 Naming latencies (RTs)

Figure 2.4 shows the pattern of RTs for correct, trimmed responses to repeated (trained) and nonrepeated (untrained) items in Experiment 2. Inspection of Figure 2.4 suggests a substantial effect of length in block 1 with, as one would expect, no difference between RTs to the items that would be repeated across the subsequent blocks of the experiment and items that would not be repeated. RTs to the repeated items followed a similar pattern across blocks 2-9 to that seen in Experiment 1, becoming faster over the early blocks then asymptoting around block 5 with a reduction in the length effect accompanying the reduction in overall RTs. Figure 2.4 suggests that RTs to untrained nonwords in block 10 were faster than RTs to untrained nonwords in block 1, but not as fast as RTs to the nonwords that were repeated between blocks 1 and 10, particularly for the longer nonwords. The analysis of the RT data was done in two parts – first an analysis of RTs to trained nonwords across blocks 1 to 10 (as in Experiment 1) and second a comparison of RTs to trained and untrained nonwords in blocks 1 and 10 only.

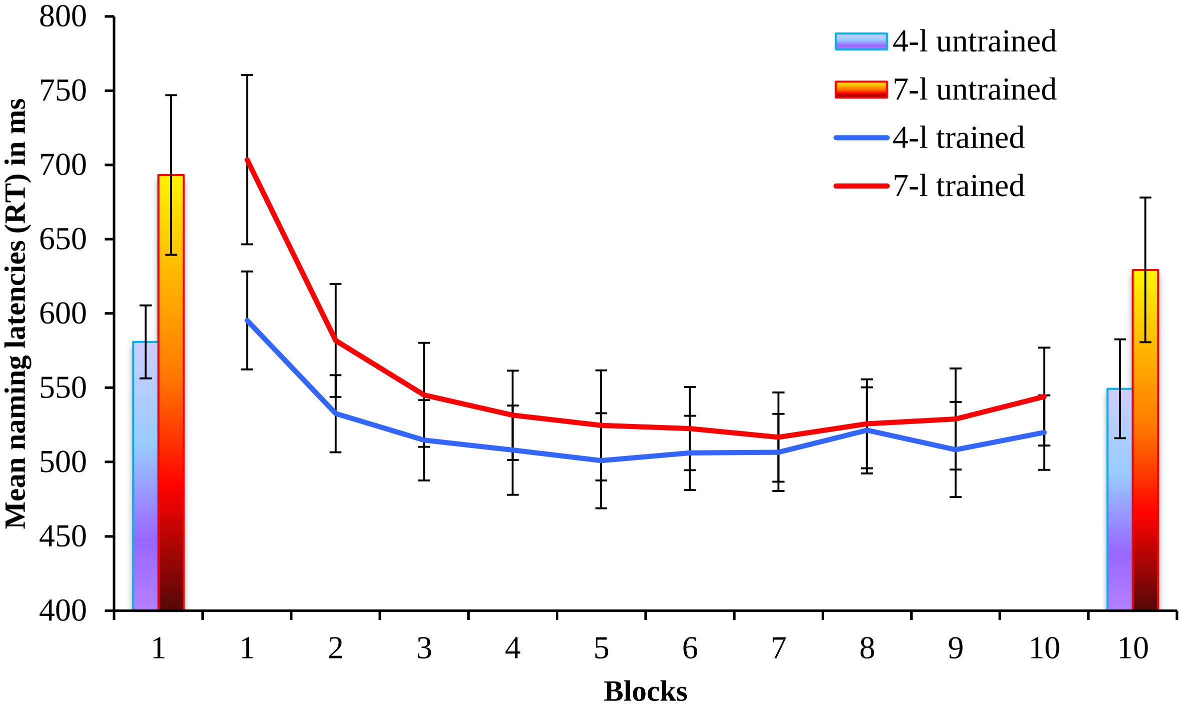

Figure 2.4. **The naming reaction times (RTs) for 4- and 7-letter nonwords in Blocks 1 to 10 in trained and untrained conditions. Error bars show 95% confidence intervals.**

2.3.2.2.1 Analysis of naming latencies (RTs) for nonwords repeated across blocks 1 to 10

Preliminary analysis of RTs to items that were repeated across blocks 1 to 10 showed no effects of sets (counterbalancing group) and no interaction of sets with the other factors. The RT data were therefore analysed in the same manner as for Experiment 1, using ANOVA with factors of Blocks (1 to 10) and Length (4 vs. 7 letters). The results are shown in section 1.1.2 of Appendix 1. As in Experiment 1, there were significant main effects of Blocks (RTs becoming faster across blocks) and Length (faster overall RTs to 4- than 7-letter nonwords) combined with a significant Blocks x Length interaction (the effect of length becoming smaller across blocks). Pairwise comparisons found significant differences between RTs to 4- and 7-letter nonwords only in blocks 1, 2, 4 and 5.

2.3.2.2.2 Analysis of RTs in blocks 1 and 10 only for trained (repeated) and untrained (non-repeated) items

Blocks 1 and 10 also included untrained items that occurred only in those blocks. The untrained items in block 10 were different from those in block 1. Preliminary analysis of RTs in blocks 1 and 10 showed no effects of sets (counterbalancing group) and no interaction of sets with the other factors. RTs to trained (repeated) and untrained (non-repeated) items in blocks 1 and 10 were therefore analysed with factors of Training (trained vs. untrained), Blocks (1 vs. 10) and Length (4 vs. 7 letters). There were significant main effects of Training (faster overall RTs to trained than untrained items), Blocks (faster overall RTs in block 10 than block 1) and Length (faster overall RTs to 4- than 7-letter nonwords). All of the two-way interactions were significant, as was the three-way Training x Blocks x Length interaction, reflecting the fact that the difference in RTs between blocks 1 and 10 was greater for trained than untrained nonwords, particularly for the longer items.

Those interactions were explored further in separate analyses of blocks 1 and 10 with Training (trained vs. untrained) and Length as factors. In block 1, the main effect of Length was significant but the main effect of Training and the Training x Length interaction were not significant (but note that at this point in the experiment, none of the items has undergone any training so effects of "Training" would not be expected).

By block 10 the trained items had been seen in each of the 9 previous blocks but the untrained items were new. In block 10 the main effects of Training (RTs faster to trained than untrained items) and Length were both significant. The Training x Length interaction was also significant, reflecting the fact that the effect of length in block 10 was 80 ms for untrained items but only 24 ms for trained items, and the fact that the difference between trained and untrained items was 29 ms for 4-letter nonwords compared with 85 ms for 7-letter nonwords. Pairwise comparisons found that the difference between RTs to 4- and 7-letter nonwords in block 10 was significant for both untrained and trained items while the difference between trained and untrained nonwords was significant for both 4- and 7-letter nonwords.

2.3.3 Discussion

RTs to nonwords that were repeated across all 10 blocks showed a similar pattern to that seen in Experiment 1, with RTs becoming faster across blocks and the effect of length diminishing. Of note is the fact that RTs to untrained nonwords also decreased between block 1 and block 10. That reduction was not, however, as great as for the trained nonwords and the effect of length in block 10 remained at 80 ms for the untrained nonwords compared with 24 ms for the trained items.

Why were RTs for nonwords seen for the first and only time in block 10 faster than RTs for nonwords seen for the first and only time in block 1? One possible explanation is based on what are termed "blocking" or "list context" effects (Lupker et al., 1997; Lupker, Kinoshita, Coltheart, & Taylor, 2003; Rastle et al., 2003). Lupker et al. (1997) and Rastle et al. (2003) compared naming latencies for high frequency words and nonwords when those two types of stimulus were either presented separately in "pure" blocks or randomly interleaved in "mixed" blocks. Naming latencies to the easier stimuli (high frequency words) were faster in pure than mixed blocks while latencies to the more difficult stimuli (nonwords) were faster in mixed than pure blocks. That is, mixing easy and difficult items had the effect of homogenising RTs to the two classes of stimuli, lowering RTs to the more difficult items while lengthening RTs to the easier items.

In block 1 of the present Experiment 2 the untrained and to-be-trained items were all new and being read aloud for the first time. By block 10, RTs to the trained nonwords had decreased considerably. The trained nonwords were now relatively easy to name, but were mixed with new, untrained nonwords that were harder to name. Under those circumstances, the influence of blocking (list context) would be expect that RTs to trained and untrained items would be homogenised, becoming shorter to the more difficult (untrained) items and longer to the easier (trained) items. In fact, the only significant consequence of mixing was the reduction in RTs to untrained set in block 10 compared with block 1. There was no apparent increase in RTs to the trained items as a consequence of being mixed with untrained items in block 10. This issue will be further explored in the General Discussion (section 2.5.1).

2.4 Experiment 3: Long-term retention of new lexical entries

After observing improvements in RT across four exposures to nonwords, Maloney et al. (2009, p. 866) remarked, "It remains to be seen how resilient these representations would be over time". A few studies have investigated possible long-term benefits of single or multiple exposures to words or nonwords.

Scarborough et al. (1977) observed a benefit for word naming latencies of a single prior naming of the same words after an interval of two days but no comparable benefit for nonword naming. One encounter with a nonword would not appear to be enough to create a representation capable of facilitating naming two days later. Salassoo et al. (1985) measured recognition thresholds for words and nonwords exposed repeatedly in 10 sessions spread over 12 days. Thresholds increased from the end of one session to the start of the next, but there was nevertheless considerable day-to-day retention of the effects of exposure for both words and nonwords. When some of the participants were re-tested a year later, thresholds for previously repeated nonwords were lower than for entirely new nonwords, indicating some retention of representations across a gap of 12 months during which the trained nonwords would not have been encountered.

Evidence of retention of representations of new written words has also been reported in studies of word learning in children. Reitsma (1983) and Share (1999) observed benefits of training on novel written words over 3-day retention intervals after the children had read the novel words some 4 to 8 times (see also Ehri & Saltmarsh, 1995, and Manis, 1985). Hogaboam and Perfetti (1978) found that training fourth grade (9- to 10-year-old) children on the spoken and written forms of novel words (nonwords) over a period of three days led to faster reading of the same items 10 weeks later. The evidence suggests, therefore, that lexical representations generated as the result of a relatively few exposures to novel words can be surprisingly resilient. That indication was tested in Experiment 3. Session 1 of Experiment 3 replicated the present Experiment 1, but with different nonwords and new participants. The participants then returned 7 days later and repeated the experiment, reading the same 4- and 7-letter nonwords in a further 10 blocks. Based on the results of Salassoo et al. (1985) and the other studies just mentioned, signs of retention of new lexical entries across the 7-day retention period, perhaps is expected, combined with some slowing of RTs at the start of the second session compared with the end of the first session.

2.4.1 Method

2.4.1.1 Participants

Forty undergraduate students of the University of York (20 male, 20 female) with a mean age of 20.6 years (range 18 - 23) took part in the experiment. All were native speakers of English with normal or corrected-to-normal vision and no history of reading or language problems. None had taken part in Experiments 1 or 2. Participants received either course credit or a small payment.

2.4.1.2 Materials

In order to ensure that the reduction of length effect was not specific to one set of nonwords, another fourteen participants (who did not participate in Experiment 1 and 2) took part in a preliminary study in which they were asked to pronounce 69 nonwords one at a time as they were shown on screen. There were 3 blocks in the pilot testing with each nonword being presented once per block. There were 2 self-paced breaks between blocks. Based on the result of the pilot testing, Set E which includes 12 pairs of nonwords which were matched on accuracy (all above 90 percent between the three blocks) and on initial letters (12 different letters to make the nonwords as different as possible) were chosen to be the experimental items in Experiment 3. The 4- and 7-letter sets were matched on initial letters and phonemes; also mean log bigram frequency. The range, mean and standard deviation of the four- and seven-letter experimental items is shown in Table 2.5. All the stimulus of Set E is shown in Appendix 6. A further 8 4-letter and 8 7-letter nonwords were created for use in the practice trials.

2.4.1.3 Procedure

Participants attended for two testing sessions, seven days apart. The Procedure for day 1 was exactly the same as for Experiment 1. Participants were asked to return 7 days later, but were not told what the second session would involve. In fact, session 2 was a repeat of session 1, including the 16 practice trials before the experimental blocks.

2.4.2 Results

Only RTs for correct responses were analysed. Naming errors, hesitations and failures to activate the voice key accounted for 306 trials (1.6% of the total). RTs shorter than 200 ms or longer than mean plus 2.5 SDs in each block for each length group were regarded as outliers and removed from the analyses of accuracy and RTs. This led to the loss of a further 313 RTs (1.6% of the total), leaving 18581 RTs (96.8% of the total) for analysis. The mean RTs (with standard deviation) in each block on each day for four- and seven-letter nonwords are shown in Table 2.6 along with the percent of correct trials in each condition.

2.4.2.1 Accuracy

Accuracy levels were high (average 96.8% correct across the two days of the experiment). Ceiling effects meant that there was no significant difference between accuracy on days 1 and 7, $W(40) = 354$, $Z = 0.51$, $p = .614$, and no overall difference in accuracy between 4- and 7-letter nonwords, $W(40) = 278$, $Z = 1.58$, $p = .15$. Figure 2.5 shows the mean accuracy for each block on Days 1 and 7.

2.4.2.2 Naming latencies (RTs)

Figure 2.6 shows the pattern of RTs for 4- and 7-letter items across blocks in day 1 and day 7. Inspection of Figure 2.6 indicates a very similar pattern on day 1 to that seen in Experiment 1. RTs then appear to have increased somewhat between the end of day 1 and the beginning of day 7, though the RTs in block 1 of day 7 were substantially faster than in block 1 of day 1 suggesting considerable retention of representations over the 7-day retention period. Figure 2.6 also indicates that by block 3 or 4 of day 7, RTs had returned to the levels seen at the end of day 1. From that point on, the difference in RTs to shorter and longer nonwords was, if anything, even less than in the later blocks of day 1.

Table 2.5. ***Mean and standard deviation of bigram frequency, neighborhood size, reading speed and accuracy of the four and seven-letter nonwords of Set E from the pilot study.***

	Nonwords	
	4-letter	**7-letter**
Bigram frequency		
Mean	2133	1908
S.D.	1045	501
Log Bigram frequency		
Mean	3.75	4.05
S.D.	0.25	0.11
Neighborhood size		
Mean	4.5	0.08
S.D.	4.46	0.29
Phonemes		
Mean	3.67	6.17
S.D.	0.49	0.83
Reading speed (in ms) in the pilot study		
Mean	546	619
S.D.	35	66
Range	481 – 626	529 – 830
Naming accuracy in the pilot study		
Mean (ppt = 14, Blocks = 3; max = 42 trials)	40	39
S.D.	1.43	1.51
Range (ppt = 14, Blocks = 3; max = 42 trials)	38 – 41	38 – 41

Note. S.D. = standard deviation; ppt = participants, max = maximum. The maximum naming accuracy is 42 as there were 14 participants in the pilot study of Experiment 3 and each of them read the nonwords aloud for 3 blocks. Thus, 14 (participants) x 3(blocks) = 42 (trials).

Table 2.6. ***Mean latencies of correct, trimmed responses, standard deviation (S.D.), and per cent correct responses for 4- and 7-letter nonwords in blocks 1 to 10 of day 1 and day 7 in Experiment 3.***

	DAY 1									
Blocks	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
4 letters										
Mean RT	597	542	522	524	510	513	507	507	505	498
S.D.	93.8	72.5	68.5	63.7	69.3	65.1	54.4	65.8	57.7	62.0
Mean Acc.	11.7	11.8	11.7	11.6	11.5	11.7	11.8	11.6	11.8	11.6
S.D.	0.78	0.38	0.58	0.63	0.78	0.51	0.49	0.67	0.49	0.71
% correct	97.7	98.5	97.1	96.7	95.4	97.7	97.9	96.7	97.9	96.7
7 letters										
Mean RT	703	585	550	540	540	526	527	516	516	510
S.D.	140.5	85.9	84.1	72.1	72.4	74.3	61.9	69.6	66.6	71.2
Mean Acc.	11.3	11.5	11.7	11.7	11.7	11.6	11.4	11.7	11.6	11.6
S.D.	0.93	0.72	0.57	0.56	0.73	0.59	0.78	0.47	0.59	0.88
% correct	94.0	95.8	97.3	97.5	97.3	96.9	95.0	97.3	96.5	96.3

	DAY 7									
Blocks	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
4 letters										
Mean RT	538	515	504	506	501	503	500	504	500	487
S.D.	87.6	86.3	65.3	72.5	73.2	80.0	80.1	76.2	83.4	74.9
Mean Acc.	11.8	11.7	11.7	11.6	11.5	11.5	11.7	11.4	11.6	11.7
S.D.	0.71	0.65	0.66	0.67	0.90	0.60	0.58	0.81	0.75	0.56
% correct	97.9	97.5	97.3	96.9	95.4	96.0	97.1	95.0	96.3	97.5
7 letters										
Mean RT	569	522	516	506	508	504	502	504	510	494
S.D.	96.7	75.6	78.0	65.1	75.2	75.8	67.0	68.1	83.1	69.8
Mean Acc.	11.7	11.7	11.6	11.7	11.6	11.6	11.6	11.6	11.6	11.6
S.D.	0.53	0.57	0.75	0.52	0.67	0.71	0.67	0.59	0.59	0.59
% correct	97.3	97.3	96.5	97.5	96.7	96.3	96.9	96.9	96.5	96.5

Note. RT = Reaction time (naming latency) in ms; S.D. = standard deviation

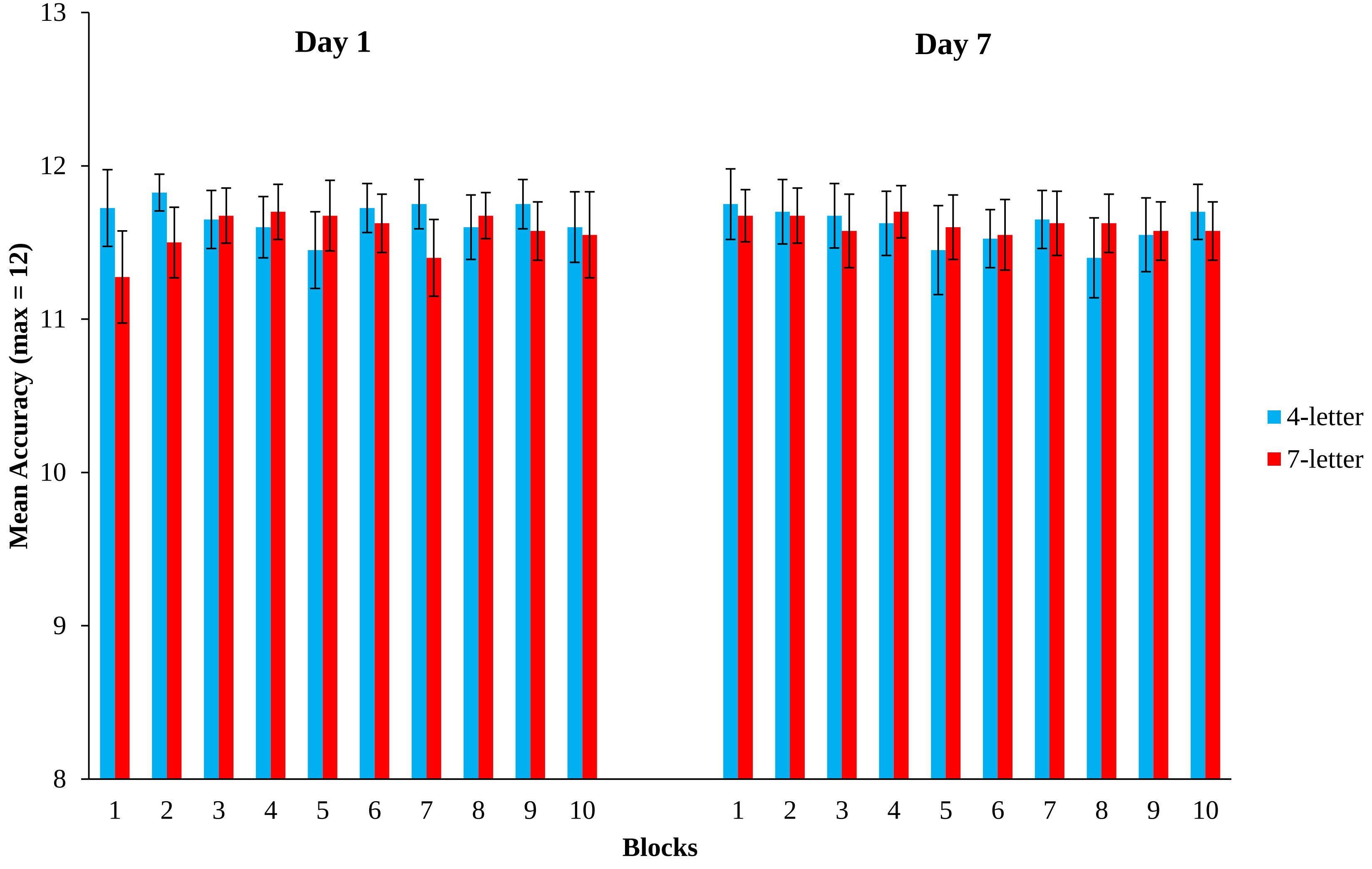

Figure 2.5. **The accuracy of naming 4- and 7-letter nonwords in Blocks 1 to 10 in Days 1 and 7. Error bars show 95% confidence intervals.**

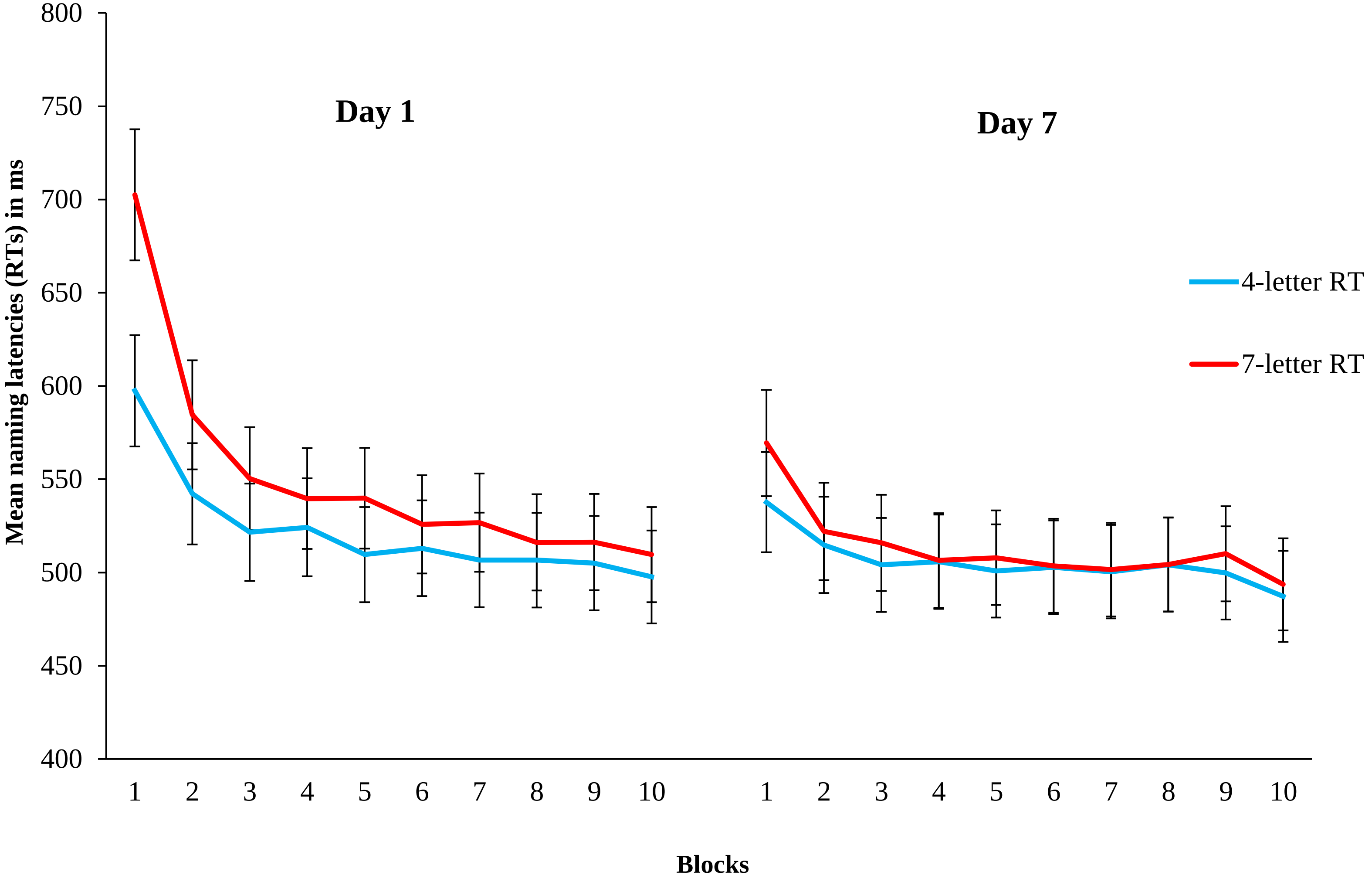

Figure 2.6. **The naming reaction times (RTs) for 4- and 7-letter nonwords in Blocks 1 to 10 in Days 1 and 7. Error bars show 95% confidence intervals.**

The RT data were first analysed across the two sessions with factors of Day (day 1 vs. day 7), Blocks and Length. There were significant main effect of Day (faster overall RTs on day 7 than day 1), Blocks (overall RTs reducing across blocks) and Length (faster overall RTs to 4- than 7-letter items). The two-way interactions between Day and Blocks, Day and Length, and Blocks and Length were all significant, as was the three-way interaction between Day, Blocks and Length, reflecting the fact that the decline in RTs across blocks and the relative change in RTs to 4- and 7-letter nonwords were greater in session 1 than in session 2. The data were analysed further through separate analyses of RTs on day 1 and day 7.

2.4.2.2.1 Day 1 RTs

Day 1 RTs were analysed with factors of Blocks and Length. As in Experiment 1, there were significant main effects of Blocks and Length accompanied by a significant Blocks x Length interaction. Pairwise comparisons of RTs to 4- and 7-letter nonwords in each block found significant differences in blocks 1 to 5 and in block 7 but not in blocks 6, 8, 9 or 10.

2.4.2.2.2 Day 7 RTs

Day 7 RTs were similarly analysed with factors of Blocks and Length. The main effects of Blocks and Length, and the Blocks x Length interaction, were significant. Pairwise comparisons of RTs to 4- and 7-letter nonwords in each block found significant differences in block 1 only. Inspection of Table 2.6 and Figure 2.6 shows that RTs to 4- and 7-letter nonwords converged numerically as well as statistically from block 4 of day 7 onwards.

2.4.2.2.3 Retention between day 1 and day 7

Retention of learning between day 1 and day 7 was assessed in an ANOVA that compared RTs in block 1 of day 7 with RTs in block 10 of day 1. The factors were Day (1 vs. 7) and Length. The effect of Day was significant (faster RTs in block 10 of day 1 than block 1 of day 7) as was the effect of Length. The interaction between Day and Length was also significant. Based on the result from the previous analyses, the effect of length was no longer significant by block 10 of day 1 but became significant again in block 1 of day 7. Comparisons between RTs at the end of day 1 and the start of day 7 found that the increase was significant for both 4- and 7-letter nonwords.

2.4.3 Discussion

The results for naming latencies in day 1 of Experiment 3 were much the same as for Experiment 1 and the trained items in Experiment 2. In the context of very high levels of accuracy, a substantial effect of length in block 1 of day 1 reduced over subsequent blocks as naming RTs decreased, becoming non-significant after approximately 6 presentations as RTs approached asymptotic levels. There was detectable slowing of RTs to both 4- and 7-letter nonwords between the end of day 1 and the start of day 7 that was followed by a re-emergence of the length effect in block 1 of day 7. But from blocks 2 and 3 of day 7 onwards, RTs were as fast as in the later blocks of day 1 and the length effect was unnoticeable numerically as well as statistically. The answer to Maloney et al.'s (2009) implied question is therefore that representations of novel words (nonwords) created by repeated exposures in a single session show considerable resilience over time, being clearly detectable in their influence on naming latencies seven days later.

2.5 General Discussion

The three experiments in the present study yielded much the same pattern of results for nonwords read aloud 10 times in 10 separate blocks within a single session. In the first block of trials, when the nonwords were read for the first time, naming latencies were slow and the difference in RTs between 4- and 7-letter nonwords was substantial. This is in line with previous reports of large effects of length on naming latencies for nonwords named only once (Juphard et al., 2004; Mason, 1978; Valdois et al., 2006; Weekes, 1997; Ziegler et al., 2001). Averaging over the present three experiments, skilled adult readers (undergraduates with English as a first language and no record of reading or language problems) read the 4- and 7-letter nonwords aloud with mean latencies of 593 ms and 693 ms respectively. An average of 33 ms per additional letter was therefore required in order to read the 7-letter nonwords compared with the 4-letter nonwords. That compares with 30 ms per additional letter in Weekes (1997), 12 ms in block 1 of Maloney et al. (2009), 34 ms in Mason (1978) and 17 ms for the English nonwords in Ziegler et al. (2001).

In each experiment, naming RTs reduced with repetition of the trained nonwords across blocks. The reduction was greater for 7-letter than 4-letter nonwords with the consequence that RTs to shorter and longer nonwords converged across repetitions. The pattern for the first four blocks was similar to that reported by Maloney et al. (2009) for nonwords of 3 to 6 letters. By block 4 in the present experiments, the mean RT for 4-letter nonwords had reduced by 83 ms compared with block 1 while the mean RT for 7-letter nonwords had diminished by 163 ms, which means that by block 4, the additional time per letter had fallen from 33 ms to 6 ms. In comparison, the mean time per letter in Maloney et al. (2009, Expt. 1), based on the comparison of RTs to 3- and 6-letter nonwords, fell across blocks from 12 ms to 0 ms. The present experiments extended training beyond four presentations to 10. Mean RTs asymptoted at around block 6. The average difference in RTs to 4- and 7-letter nonwords blocks 6 to 10 of the present Experiments 1 to 3 was stable at around 14 ms, giving a mean time per additional letter of just 5 ms.

The evidence of the strong claim that length effects are completely eliminated by 5 or 10 exposures to novel words within a single session is not advocated in this study. Given the reports in the literature of effects of length on naming latencies for real words, especially for low-frequency words, the result of the study assert that greater improvement across presentations in RTs to longer than shorter nonwords means that length effects are greatly reduced by repeated exposures within a session.

2.5.1 The possible contribution of blocking / list context effects

What contribution, if any, might blocking / list context effects make to the pattern of results observed in the present experiments? In the Experiment 2, untrained nonwords were read more quickly when they were interleaved with trained nonwords in block 10 than when they appeared in block 1 with nonwords that were also being read for the first time in block 1. The untrained nonwords in block 10 were not read as quickly as the nonwords that had received training in the previous 9 blocks, and the effect of length was greater in the untrained than the trained nonwords, but their naming RTs benefited from being mixed with the trained nonwords.

In the Discussion section of the present Experiment 2, it is noted that the observation by Lupker et al. (1997) and Rastle et al. (2003) that high-frequency words are named more slowly when mixed with nonwords than when presented in unmixed ("pure") blocks of trials while nonwords showed the opposite pattern, being named more slowly when presented on their own in pure blocks than when interleaved with high-frequency words in mixed blocks of trials. Lupker et al. (1997) and Rastle et al. (2003) argued that participants set a criterion for the speed of responding to stimuli in a block based on the blend of easy or difficult items within the block. When the items are all easy (e.g., pure blocks of high-frequency words) the criterion will be relatively short and RTs consequently faster. When the items are all difficult (e.g., pure blocks of nonwords) the criterion will be relatively long and RTs slower. When the items are a mixture of easy and difficult, a criterion will be set that is somewhere between in the middle resulting in a homogenization of RTs to easier and more difficult items.

Taylor and Lupker (2001) went on to show that criterion shifts (if that is what they are) can be adjusted on a trial-by-trial basis rather than across a sequence of trials so that the naming latency for a particular item in a sequence will be influenced by the ease or difficulty of naming the preceding item. Their Experiment 3 investigated the effects of blocking and preceding trials using "fast" (easy) and "slow" (difficult) nonwords (categorised on the basis of their RTs in an earlier experiment). Easy nonwords were named faster in pure than mixed blocks. Within the mixed blocks, the easy nonwords were named faster following other easy nonwords than following difficult nonwords. In contrast, RTs for difficult nonwords were not significantly different between pure and mixed blocks, and within the mixed blocks there was only a trend for RTs to be faster following easy nonwords than following other difficult nonwords. Reynolds, Mulatti, and Besner (2012) obtained a similar pattern of results using a paradigm more associated with task switching than blocking effects. Easy and difficult nonwords were presented in a predictable AABB order rather than in a random order as in the mixed conditions of blocking experiments. RTs to easy nonwords were faster following other easy nonwords than following difficult nonwords but RTs to the difficult nonwords were not significantly affected by switching.

Reynolds et al. (2012) did not associate their findings directly with those of Taylor and Lupker (2001) but their results are clearly similar in finding RTs to easy nonwords to be more affected by list context than RTs to difficult nonwords. That was not the pattern seen in the present Experiment 2 where untrained (difficult) nonwords was benefited in block 10 from being mixed with trained (easy) nonwords but RTs to the trained (easy) nonwords were barely affected (if at all) by being mixed with new, difficult nonwords.

The original study by Weekes (1997) interleaved high frequency words, low frequency words and nonwords of varying lengths. List context effects should mean that RT differences between conditions were reduced as a result of homogenisation. That could apply to short and long nonwords within an experiment as much as to high frequency words, low frequency words and nonwords. In Maloney et al. (2009) and the present experiments, the use of mixed lists of nonwords of different lengths should mean that RTs to easier (shorter) nonwords are slowed by the presence of harder (longer) nonwords, and conversely. As learning continues and all the nonwords become easier, the criterion for response production should be revised down, resulting in a general reduction in RTs. The results of the present Experiment 2 show, however, that this is not the whole story. RTs in block 10 of that experiment remained slower to untrained than to trained nonwords, and the convergence of RTs to shorter and longer nonwords was much more apparent for the trained than the untrained items. That said, and despite discrepancies between the present results and those of Taylor and Lupker (2001) and Reynolds et al. (2012) that need to be explained, the list context (blocking) effects may play a part in generating the overall pattern of effects seen in this and similar studies.

2.5.2 Creation of lexical representations and the modulation of the length effect

What underlies the reduction in the effect of length across repeated exposures to novel words (nonwords)? Within the framework of dual-route models, differences in length effects have been regarded as core phenomena requiring explanation (Coltheart et al., 2001). That explanation involves proposing that as novel words become familiar, representations are created for those words in both the orthographic input lexicon and the phonological output lexicon. This allows processing to switch to occur from a serial (and therefore length-sensitive) nonlexical route to a lexical route in which the component letters of words are processed in parallel.

When the nonword's letter string is first presented to the DRC model in Block 1 of Day 1, the model had to pronounce these nonwords through applying grapheme-phoneme rule system and pronounce the nonwords through administrating the non-lexical route. As the nonwords are processed in a sequential, left-to-right form, a robust length effect can be found in this stage. For example, the nonword 'brup' has to go through the process of b→/b/, r→/r/, u→ /ʌ/, p→ /p/. Since there are more letters in a 7-letter nonword, it will take the model longer to pronounce a 7-letter compared to 4-letter nonword. At this stage, naming latency is around 600ms for 4-letter nonwords, and 700 ms for 7-letter nonwords. Referring back to Weekes (1997) study, this is very similar to the naming latency of nonwords.

Moving forward to Blocks 2 to 7 in Day 1, though the model was still partly processes the nonwords in a serial way, it was also creating lexical entries in the orthographic input and phonological output lexicon. As the lexical route processes relatively slowly, a small but significant contribution from the non-lexical route can still be observed. At this stage, the naming latency of 4-letter nonwords was around 520 ms and 7-letter nonwords was around 550 ms. Referring back to Weekes (1997) study, this is very similar to the naming latency of low-frequency words.

When it comes to Block 8 to 10 of Day 1, the naming performance was fully dominated by the lexical route. Though the non-lexical route cannot stop its' contribution towards the naming tasks, given that the lexical route operates very quickly, verbal response is delivered by the lexical route before the non-lexical route is able to produce any responses. Thus, no detectable contribution from the non-lexical route is observed. After approximately six exposures, unitization is fully completed by this stage and it is the result of the formation of lexical entry in the orthography input and phonological output lexicons. This can be indexed by the reduction in the magnitude of the letter length effect. At this stage, the naming latency of both 4- and 7-letter nonwords were both around 500 ms. Referring back to Weekes (1997) study, this resembles the naming RTs of high-frequency words.

2.5.3 Length and neighbourhoods

Shorter nonwords typically resemble several other words while longer nonwords tend to be more distinctive in their appearance. Resemblance between words, or between nonwords and words, is conventionally measured in terms of other words that can be generated by changing single letters or phonemes in a particular word or nonword. "Orthographic neighbours" are other words that can be generated by changing a single letter in a word or nonword while "phonological neighbours" are other words that can be generated by changing a single phoneme in a word or nonword: *tough* is both an orthographic and a phonological neighbour of *rough*, *dough* is an orthographic but not a phonological neighbour and *huff* is a phonological but not an orthographic neighbour. The number of words that can be generated by changing single letters in a word or nonword is known as "orthographic N" while the number of words that can be generated by changing single phonemes is known as "phonological N".

Shorter words and nonwords tend to have more neighbours than longer words and nonwords which means that length and N are naturally correlated. It is possible, therefore, that some part of the effects attributed here to variations in letter length and in fact attributable to variations in N. Balota et al. (2004), Cortese and Khanna (2007) and Morrison and Ellis (2000) found independent effects of both letter length and orthographic N on word naming latencies, implying that both factors contribute to determining naming latencies to mixed sets of words. No attempt was made in those studies to distinguish orthographic and phonological aspects of length or neighbourhood size. Yap and Balota (2009) found effects of both orthographic and phonological N on word naming latencies plus a separate effect of letter length. In Cortese and Schock's (2013) analysis of word naming the effects of letter length and orthographic N were both significant but the effect of phonological N was not. Using sets of words matched on N, Lavidor and Ellis (2002) found an effect of letter length on lexical decision RTs for words presented in the left visual field (LVF), but not for words presented centrally or in the right visual field (RVF), indicating that the differential effects of length in the LVF and RVF for lexical decision do not reduce to differences in N.

Effects of letter length on visual word recognition do not appear, therefore, to reduce completely to effects of N, but variation in orthographic N could still contribute to the pattern of results seen in the present experiments. For that to be a factor, the effect of N on naming latencies should be greater for words than nonwords, greater for high than low frequency words, and greater for words in the LVF than the RVF. Somewhat surprisingly (given the large amount of research devoted to effects of N on word recognition), there appears to be only one study that has compared the effects of N on naming latencies for words and nonwords. That study (Perea & Carreiras, 1998) found similar effects of N on naming speeds for Spanish words and nonwords which is not the pattern we would expect if variation in N contributes to the length by lexicality interaction. On the other hand, Andrews (1989; 1992) found larger effects of N on naming latencies for low than high frequency words which mirrors the larger effects of length for low than high frequency words. Evidence for parallel effects of N and length in the two visual fields was presented by Lavidor and Ellis (2002) who found larger effects of neighbourhood size in the LVF than the RVF (though in lexical decision rather than word naming). Further research is needed to clarify the relationship between the effects of length and N, particularly with regard to Perea and Carreiras’s (1996) report of comparable effects of N on word and nonword naming in Spanish. The findings for high vs. low frequency words and LVF vs. RVF presentations are more compatible with the notion that length and N have similar but independent effects on word and nonword naming.

2.5.4 Retention versus decay of lexical representations

In the present Experiment 3, naming RTs increased between the end of the first testing session and the start of the second session 7 days later. But after just two or three presentations in that second session, RTs had decreased back to the level seen at the end of the first session and the convergence of RTs to shorter and longer nonwords was virtually achieved. A similar result was obtained by Salasoo et al. (1985) who found that in the earlier blocks of training, recognition thresholds increased between the end of one session and the start of the next, but to a level below that seen at the start of the previous session. With further presentations, thresholds fell until they eventually asymptoted. Taken together, these observations suggest that in the absence of exposure to the novel words, representations may undergo a small amount of decay or forgetting between the end of one session and the start of the next, combined with a considerable degree of retention that allows the representations to strengthen further after just a few presentations.

The DRC model of Coltheart et al. (2001) does not learn through experience; neither does it forget. It can be programmed to simulate different degrees of learning, but lacks the ability to create new lexical entries in response to training and there is nothing in the DRC model analogous to loss of representational integrity through decay or interference. Participants in the present Experiment 3 will not have encountered the trained nonwords in the interval between the end of the first training session on day 1 and the start of the second session on day 7. They will, however, have encountered a great many familiar words, creating the circumstances under which experience with those words could have interfered with the newly-established representations of the experimental nonwords. That could account for the decline in performance between sessions that is visible in the results of the present Experiment 3 and also in the results of Salasoo et al. (1985).

Loss of representational integrity as a result of interference from other words might also account for finding that length effects are greater for low- than high-frequency words. High- frequency words like *bed* or *cut* are likely to be encountered often enough to resist interference from other words. In contrast, low-frequency words like *wig* or *grid* may only be encountered in written form a few times a year. Interference from other words may disrupt their representations sufficiently to allow length effects to appear.

By their very nature, frequency effects imply that regular encounters with words makes representations more efficient and capable of being activated more rapidly. Set against that we have evidence that representations created as the result of relatively few encounters with a new word can survive over long periods when they are not activated (e.g., Salasoo et al., 1985, and the present Experiment 3). What might prevent the representations of novel words that are not regularly refreshed by additional encounters from suffering catastrophic interference and loss? One possible mechanism is provided by the "complementary learning systems" approach to learning presented by McClelland, McNaughton, and O'Reilly (1995) and applied to word learning by Davis and Gaskell (2009). The complementary learning systems approach proposes that when new connections must be created between representations in different parts of the brain (e.g., the orthographic and phonological representations of novel words), the hippocampus and associated cortex is initially involved in building those connections. Over time, and as a result of consolidation processes that may be facilitated by sleep (e.g. Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010), those connections are established at a purely cortical level, freeing the hippocampus for new learning. O’Reilly, Bhattacharyya, Howard, and Ketz (2011) suggested that consolidation and transfer of information to the cortex helps protect against interference. A hallmark of the transfer from hippocampal to cortical connections is the emergence of competition effects between newly-learned words and established vocabulary (e.g. Henderson, Weighall, Brown, & Gaskell, 2013). If so, then under the conditions of the present experiments, competition between novel written words and established words in the lexicon of the sort reported by Bowers et al. (2005) should be observed after a period of consolidation (e.g., session 2 of the present Experiment 3), but not within the initial learning session.

2.6 Conclusions

The three experiments reported here found that repeatedly presenting novel words (nonwords) to be read aloud as quickly as possible results in a reduction of naming latencies and a decrease in the impact of length. The dual-route approach attributes faster naming and parallelisation of processing to word-specific learning within the lexical system. The increased fluency of naming that accompanies learning may itself contribute to the facilitation of naming through blocking (list context) effects, but Experiment 2 shows that such effects cannot account for the full facilitation of naming speeds or the convergence of RTs to shorter and longer items.

The results of all three experiments suggest that an average of four to six exposures to a novel word is sufficient for skilled readers to create lexical representations which capture the process of word naming. That estimate of the number of exposures required for the establishment of lexical representations agrees with previous studies of visual word learning in both children and adults (Ehri & Saltmarsh, 1995; Feustel, Shiffrin, & Salasoo, 1983a; Hershenson & Haber, 1965; Hogaboam & Perfetti, 1978; Maloney et al., 2009; Manis, 1985; Reitsma, 1983; Salasoo et al., 1985; Share, 1999; Solomon & Postman, 1952). Once established, the novel lexical representations prove to be remarkably resistant to decay or interference, even when they are not refreshed by further exposures to the novel words.

There are, as always, issues remaining to be resolved, one of which is the extent to which differences in neighbourhood density between shorter and longer nonwords contribute to the effects observed here and elsewhere. Yet, the paradigm developed here has considerable potential as a tool for investigating visual word learning. One application would be to study word learning in different groups of readers (e.g., dyslexics or second language learners whose first language is or is not alphabetic, see Chapters 5 and 6). By varying the nature of the training provided to participants, it should also be possible to investigate the relative contributions of orthographic and phonological learning to the effects observed here (cf. Hogaboam & Perfetti, 1978; Maloney et al., 2009; McKay, Davis, Savage, & Castles, 2008; Reitsma, 1983; see Chapter 3), and the possible additional impact of associating meanings with the novel words, as happens in natural language learning (cf. McKague, Pratt, & Johnston, 2001).

3 The role of phonology in visual word learning

3.1 Introduction

Orthographic learning continues across the lifespan --- skilled adult readers persist to encounter new words in both spoken and written domains, and these have to be amalgamated into the existing lexicon without compromising the accurate and efficient recognition of words that are already familiar (e.g. Grossberg & Stone, 1986). Despite this fact, surprisingly little research has investigated on-going learning in the visual word recognition system.

This chapter utilizes the learning paradigm that was introduced in Chapter 2 to investigate orthographic learning in skilled adult readers, integrating theories of reading development with the skilled reading literature. The 'item-based' account of lexical acquisition put forward by theorists including Share (1995, 1999, 2004), Ehri (1989, 1992) and Perfetti (1992, Perfetti & Hart, 2001) is elaborated to address the role of phonology in orthographic learning. Share's theory of phonological recoding (print-to-sound translation) as a lifelong self-teaching process is extended to explore the potential role of feedback from phonology in the process of orthographic lexical acquisition of new words --- a process that McKague et al. (2008) referred to as *orthographic recoding*.

3.1.1 Lexical equality hypothesis

A fundamental prerequisite for the high level of proficiency in reading and spelling achieved by educated adults is a well-established memory representation for each word in one's lexicon. This is word specific knowledge that can be segregated into orthographic, phonological and semantic components (Perfetti, 1992). According to Perfetti, a high-quality lexical representation is complete and accurate in all three components with efficient links between the components. As the semantic component has been found to only affect irregular word learning (Nation & Cocksey, 2009), the three experiments that are included in this Chapter do not include the semantic component.

This efficient linkage among the three components is unlikely to be achieved in a single encounter with a printed word. According to item-based accounts of orthographic learning, each word item must undergo an individual process of specification, in which the precise and integrated connections are formed. Thus, even for the skilled reader, the orthographic lexicon may contain representations that vary in their degree of specification (Perfetti & Hart, 2001). Perfetti defines moderately specified representations as that skilled readers are aware of the letters that are involved in a target word, but they are not certain of the specific position of each single letter.

According to item-based developmental theories, the mapping of orthography to phonology is the process that supports the development of precisely and fully-specified orthographic representations. It is the sequential processing demanded by the mapping of orthography and phonology that assures both precision in encoding the letter sequence and the overlapping levels of connections between orthography and phonology (Landi, Perfetti, Bolger, Dunlap, & Foorman, 2006). In favour of item-based accounts, several studies demonstrate forming links between orthography and phonology helps the development of orthographic representations for both beginning readers (Cunningham, 2006, Manis, 1985; Share 1999, 2004) and adults (Brooks, 1977; Sandak et al., 2004).

Few would disagree with the claim that building linkage between orthography and phonology is necessary for the acquisition of orthographic knowledge of the item. Yet, current research in the field has mainly focused on the unidirectional flow of information from orthography to phonology which makes it insufficient to explain the process of how establishing orthographic representations come to be as strongly determined by phonemic factors as is conveyed by item-based theorists. Ehri (1992) suggests that ‘Orthographic representations are paved with phonological information.’ This implies that phonological representations would inevitable be activated in the orthographic learning process. The hypothesis tested in the present chapter is that feedback from phonology to orthography, or *orthographic recoding* (McKague et al., 2008), plays a role in the process of orthographic learning in an implicit learning paradigm.

3.1.2 Orthographic learning in children

Reitsma (1983, experiment 1) explored the mechanism of orthographic and phonological learning by asking third grade primary school children to perform a lexical decision task. In the training session, six words were introduced to each child in association with pictures: half were fictitious animals and half were imaginary fruit. Children had to learn to make a categorical decision upon presenting the novel words (animal or fruit). During the training, half of the novel words were presented only in the auditory domain (A) and half were presented visually (V) as well. As shown in Figure 3.1, it only took participants four trials in order to learn the novel words that were presented in the visual and auditory domain. Children were able to acquire the knowledge of the novel words that were presented phonologically, yet learning from the visual domain was better in the first three blocks. The process of phonological word learning in children was then further developed by Share (1995, 1999, 2004), Cunningham et al. (2002, 2006), Ehri (1992), Kyte and Johnson (2006) (see Chapter 1).

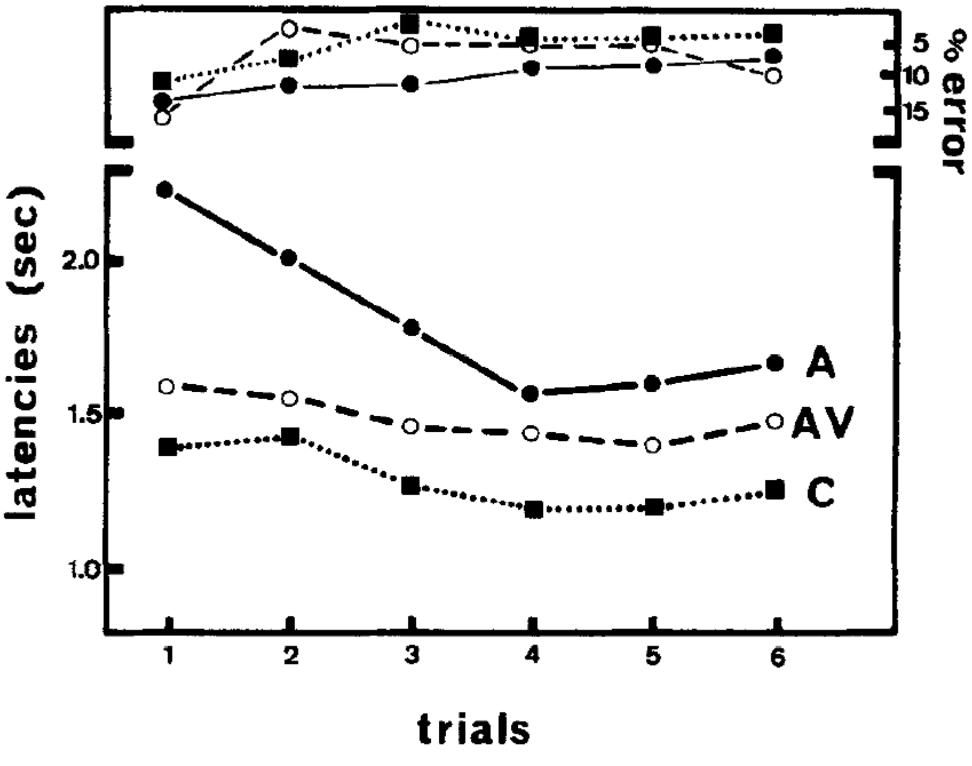

Figure 3.1. **Mean naming latency (in seconds) for common words (C), auditorily learned pseudowords (AV) and visually learned pseudowords (A) (taken from Manis 1985).**

As mentioned in Chapter 1 (section 1.5), Ricketts, Bishop and Nation (2009) investigated the integration of orthography and phonology by exploring whether exposure to orthography facilitates oral vocabulary learning. Children showed robust learning for novel spelling patterns after incidental exposure to orthography. Furthermore, there was stronger learning for nonword-referent pairings trained with orthography. This is consistent with previous studies that show that children are more likely to learn phonological forms when they are presented with orthographic information (e.g. Ehri & Wilce, 1979; Hu, 2008; Hulme et al., 2007; Reitsma, 1983). Furthermore, the authors interpret this finding as demonstrating that learning is improved for word representations that include orthographic, phonological, and semantic information (Perfetti & Hart, 2002).

A similar finding is reached by Rosenthal and Ehri (2008). They asked fifth graders to learn 10 words; for example, vibrissae (the whiskers on a cat) and tamarack (a huge tree). In the experiment, the words were pronounced, defined, embedded in sentences, and depicted in drawings on flash cards. Children were given several practice trials to learn the pronunciations and meanings of the words. On each trial they were prompted to recall either the pronunciation or the meaning of each word. In one condition, spellings appeared on the cards during study and feedback periods but not when children recalled the words. In the control condition, the same procedures were followed except that students were not shown spellings. Instead, they pronounced the words a few times.

Figure *3.2* shows the result of how well the children recalled pronunciations of the words across learning trials when spellings had or had not been seen. Results indicated that for both high and low level readers, their recalled accuracy in trial 5 was 30% worse in the spelling-not-seen compared to the spelling-seen condition. This illustrated that children’s learning of the pronunciation of novel words was hindered by not seeing the orthography of the stimuli.

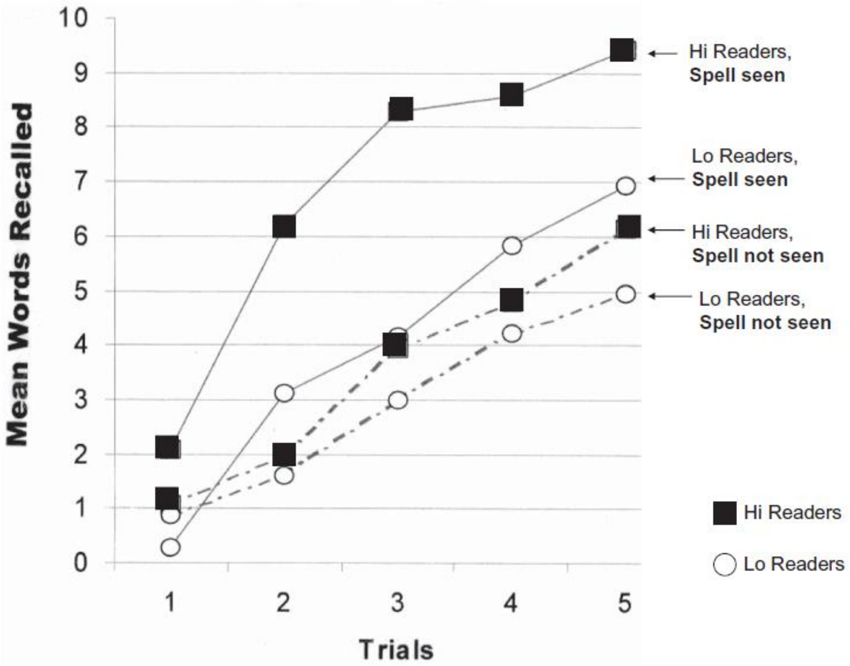

Figure 3.2. **Mean number of pronunciations (10 maximum) recalled over five learning trials by higher and lower ability fifth-grade readers in the spelling seen and spelling not seen conditions. (taken from Rosenthal and Ehri. 2008).**

3.1.3 Orthographic learning in adults

This section will focus on the process of word acquisition in adults. In Experiment Two of Maloney et al.’s (2009) study, participants were separated into two groups: a case decision group and a reading aloud group. In the four blocks of the case decision task, participants were asked to verbally identify the case in which a letter string was presented by responding ‘upper’ or ‘lower’ aloud. In the four blocks of the reading aloud task, participants were asked to read the letter string aloud. In the fifth block (the test block) all participants were instructed to read aloud the letter string. Figure 3.3 shows the main result of the reading aloud group. Their result found there was a significant Block x Letter Length interaction in the reading aloud group which indicated that the magnitude of the letter length effect decreased significantly across blocks.

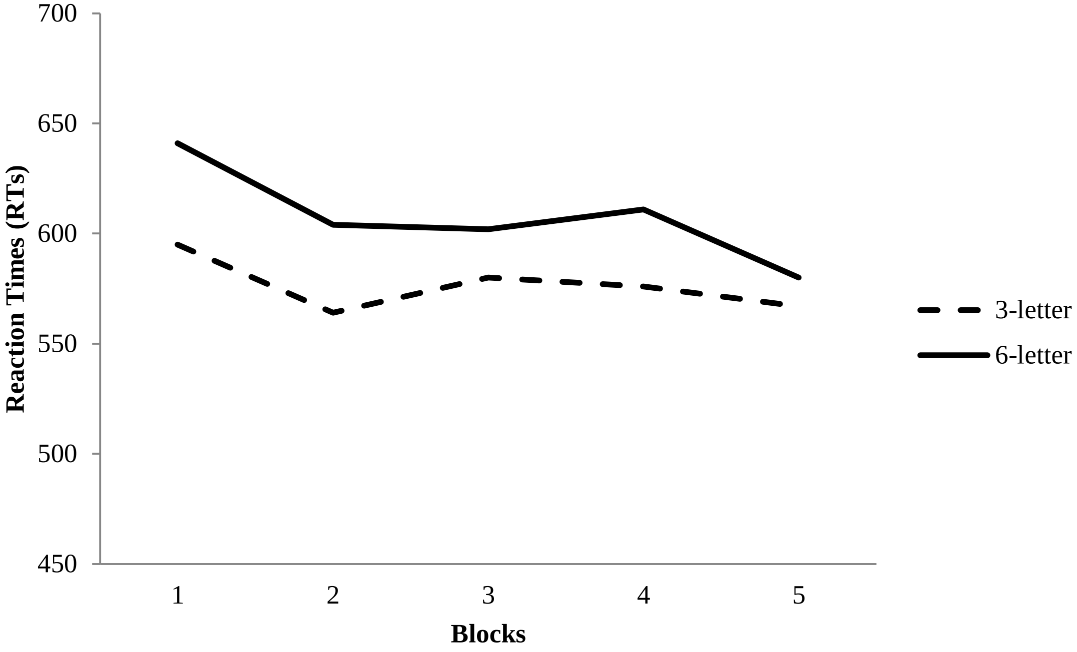

Figure 3.3. **The naming RTs of the reading aloud group in Maloney et al. (2009).**

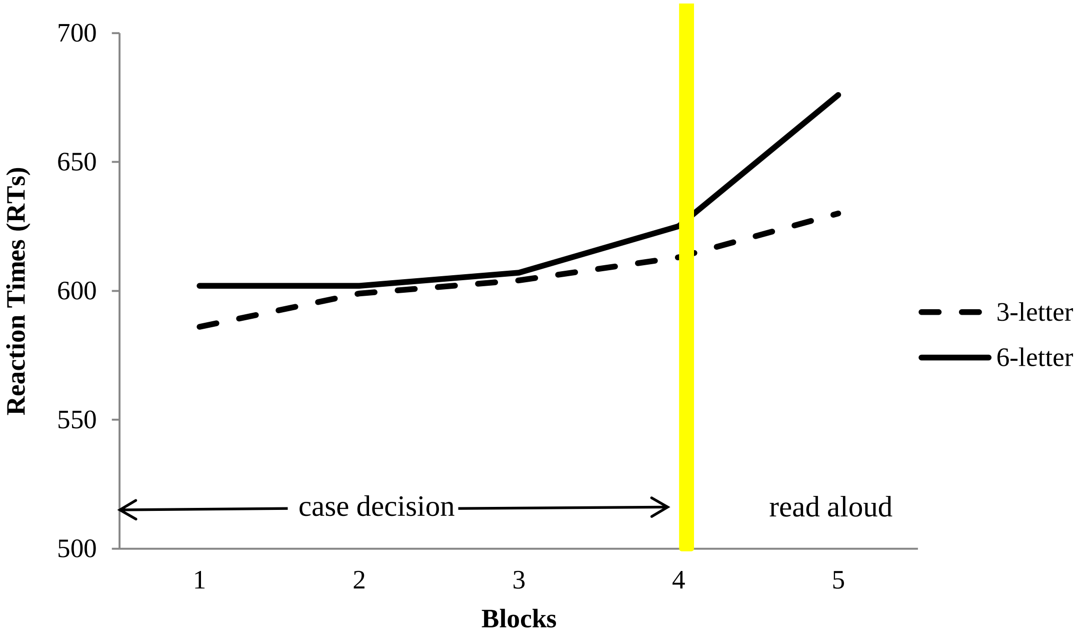

Figure 3.4. **The mean response times of the case decision group in Maloney et al. (2009) study.**

Figure 3.4 shows the main result of the case decision group in the Maloney et al. (2009) study. The case decision task was chosen because the researchers claimed the task did not require the generation of a phonological code and could provide a control group for the reading aloud group. The result showed there was a main effect of blocks for the case decision group, in which the case decisions were made more slowly in Block 4 than in Block 1. The letter length effect was only significant in the by-item analysis. There was no improvement in RTs in the case decision condition. The naming RTs of Block 5 (reading aloud block) in the case decision group was also slower than Block 1 of the reading aloud group. Given the result of the case decision group, Maloney and colleagues proposed the 'automatic' generation of a phonological code was not sufficient to form a lexical representation. Rather, they suggested it may be the 'explicit' generation of a phonological code is required in order to form represention in the mental lexicon. Yet, there is a methodological flaw that may have blurred the result of the study. Participants in the case decision task had to experience a switch from case decision task to reading aloud in block 5 whereas participants in the read aloud group did not experience any task switching. Thus, it is not certain whether the difference of the reaction times in blocks 4 to 5 in the case decision condition was due to the training or the task switching effect.

As mentioned in Chapter 1 (section 1.6), acknowledging that there is rapid learning in orthographic and spoken word learning in adults, Chalmers and Burt (2008) took a further step to understand the role of phonological encoding skills in orthographic learning. The results showed that the provision of either phonological or semantic information during training improved spelling recognition. A similar result was obtained and extended by Nelson, Balass, and Perfetti (2005), Taylor, Plunkett, and Nation (2011), and Sandak et al. (2004). There is a key point of Chalmers and Burt (2008) study. It is true that the combined (phonological learning + orthographic learning) is significantly better than the condition that was solely trained on the orthography of the novel words (counting the consonant that's in a novel word), but just by training participants on orthography of the word is significant enough to promote word learning. Spelling recognition accuracy in the orthography condition was 73% versus 81% in comparison to the combined learning condition. As Chalmers and Burt (2008) suggested, ‘It was important for the assessment of encoding effects that participants were not informed about the nature of the subsequent test. Participants expecting a spelling test on the items may have changed their encoding strategies.’ This is the main reason that Experiment 4, 5 and 6 below were specially designed in a way that participants were not aware that they were going to see new filler items in the final block while they were in the learning phase.

Chalmers and Burt (2008) study has implications for the nature of orthographic learning and the individual skills that support it. First, in terms of encoding information and strategies, the results reinforce the importance of relating orthography to phonology when learning to spell new words. In this respect, the study is consistent with the results of training studies in children that have found beneficial effects of item pronunciation during study (e.g. Kyte & Johnson, 2006). In line with the self-teaching hypothesis, it appears that linking phonology to orthography facilitates attention to and retention of an unfamiliar word’s letter sequence (Share, 2004).

Secondly, the result of Chalmers and Burt (2008) also supports the notion that phonological encoding skills play a role in orthographic learning in adults. In the developmental literature, it is commonly held that phonological coding is important early in reading but not when reading becomes highly fluent (e.g. in older children and adults, Frith, 1986). Dual process theory holds there are two independent processes for reading single words: an indirect process which relies on grapheme-phoneme-correspondence rules and a quicker direct process which accesses word-specific knowledge from orthography. Normal readers are able to use both processes, but normal development is seen as a progression from an early reliance on the slower indirect process, which requires phonological coding, to later reliance on the direct process. Pennington, Lefly, Van Orden, Bookman, and Smith (1987) have referred this assumption as ‘the phonological bypass hypothesis’ since it assumes phonological coding is eventually bypassed in normal reading development. Pennington et al (1987) had disputed several predictions that they derived from it. For instance, they demonstrated that the phonological coding skill also predicts huge amount of variances in adults' reading measures. The results of Chalmers and Burt (2008), together with those of Sandak et al. (2004), challenge this view by providing evidence that phonological coding may play an important role in orthographic learning in adults. In sum, although adults may show in orthographic learning more sophisticated knowledge about English orthography and morphology than children do, there is little reason to suppose that the processes of orthographic learning are fundamentally different in early readers and adults.

Recognizing that the role of phonology is salient in word learning, McKague et al. (2008) explored whether a briefly formed orthographic representations of the novel words would be activated when participants received phonological training. Eight-four participants (42 in each condition) were separated into oral (n = 44) and visual (n = 40) instantiation training groups. In the oral instantiation training conditions, participants learned 32 rare English words by watching a video of a narrator talking about them on the computer screen. In the visual instantiation training, the procedure was identical to the oral instantiation training except that participants read each of the passages silently. The participants in both the oral and visual instantiation conditions completed a visual lexical decision task at the end of the instantiation training session. Participants had to indicate whether the stimuli were words (including the instantiated words) or nonsense words. Each trial of the visual lexical decision task commenced with the display of the lower-case priming stimulus before the target stimuli was shown: 1) an identity prime (lerse/LERSE), 2) a consonant-preserving form prime (a single vowel letter was altered; lorse/LERSE), 3) a vowel preserving form prime (a single consonant letter was altered; lerve/LERSE) and an all-letter-different control prime (spolt/LERSE). The brevity of the prime meant that participants were rarely able to report it, and it was not open to slow decoding or strategic influences. Thus, any facilitation produced by the prime is assumed to reflect the fact that the prime has rapidly and automatically activated the orthographic representation for the target word (Forster, Mohan, Hector, Kinoshita, & Lupker, 2003).

The result showed that orally instantiated novel words preceded by the identity prime were recognised significantly faster than those preceded by the consonant-preserving form prime. This meant that there was an inhibitory effect of feedback inconsistency for orally instantiated novel words. Furthermore, orally instantiated novel words received significantly more facilitation from consonant-preserving form primes than from vowel-preserving form primes. The result support the notion that there is a reciprocal bidirectional connection that forms between orthography and phonology in the process of learning to read and write that enables skilled readers to automatically recode novel phonological inputs into orthographic codes before printed exposure is appealing at an intuitive level. This study had extended and replicated the study of Johnston et al. (2004). Similar results were obtained by Rastle et al. (2011) and Bürki et al. (2012). The experiments in this chapter were not designed to investigate the feedback consistency effect. Yet, the fact that previous literature showed the orthographic representations were activated automatically in phonological training meant that this chapter has to control for automatic activation of orthographic codes when participants were in the phonological learning conditions in Experiment 6.

3.1.4 Predictors of word reading in adults

As mentioned in section 1.4 of Chapter 1, a good amount of developmental studies illustrate that phonological skill is not only important in word learning skills in children (e.g. Bowyer-Crane et al., 2008), but also in adults (Young et al., 2002). Ricketts, Bishop & Nation, 2009) showed nonword reading skills (TOWRE PDE) and word reading skills (TOWRE SWR) significantly correlated with the ability of spelling. The result also suggested that more advanced readers showed more benefit from orthography in the training phase. Hulme et al. (2007) obtained a similar result that phoneme deletion is a significant predictor in nonword reading in children. As phonological skill is not the only predictor that can explain all the variance in word learning, more studies have now focused on other factors that are equally important in word learning. This includes factors like vocabulary skills (e.g. Braze, Tabor, Shankweiler, & Mencl, 2007) and rapid digit naming (RAN; Wolf, 1997).

3.1.5 Orthographic consistency effect

As mentioned in section 1.2 of Chapter 1, the role of feedback from phonology to orthography in visual word recognition is controversial, especially for novel words (e.g. Peereman, Content, & Bonin, 1998). The feedback consistency effect occurs when lexical decision or naming reaction times are slower for words whose pronunciations can be spelled in several ways. Stone, Vanhoy, and Orden (1997) were the first study to report an inhibitory effect of feedback inconsistency in visual word recognition using the lexical decision task. The effect was then found in a speeded naming task as well (Ziegler, Montant, & Jacobs, 1997). The explanation for the feedback consistency effect is that it demonstrates automatic feedback from the activated phonological code to the orthographic level such that potential spelling representations compete.

It is informative to note that most of the studies reporting feedback consistency effects have utilized items of low frequency – often between 1 and 10 occurrences per million (e.g. Stone et al., 1997; Ziegler & Ferrand, 1998; Ziegler et al., 1997). It will be beneficial to understand whether this process could be extended to learning novel inputs. McKague et al. (2008) suggested that the feedback consistency effect may be an essential step for learning new items. Phonological feedback can help to refine the perception of the orthographic code, and assures the encoding of the correct sequence of letters is distinguishable from other possible spellings of the computed phonology. Ventura, Morais, Pattamadilok and Kolinsky (2004) elaborated the idea that the less precise phonological code would benefit from being grounded in the visual orthographic code.

3.2 Experiment 4: The role of phonology in orthographic learning (within subject design)

This chapter reports three experiments investigating the role of phonology in visual word learning in skilled, adult readers of English. Experiment 4 represents an extension of Maloney et al. (2009, experiment 2). All participants read aloud the novel words in Block 1, there were two types of training after Block 1 and participants had to read aloud novel words again in the final block. There were two types of training, hear-and-repeat and read aloud, all participants received both training with half the participants going through read aloud training first before they received the hear-and-repeat training. On the basis of developmental study (e.g. Reitsma, 1983 and McKague et al, 2008), it is expected to see that though hear-and-repeat training would be sufficient to build certain representation in the orthographic lexicon, there would be a greater improvement for participants who are trained in the read aloud condition.

It is hypothesized that RTs would reduce across blocks in both conditions. Yet, the RTs to shorter and longer items would converge more in the read aloud condition compared to the phonological training condition. Furthermore, given that previous studies (e.g. Weekes, 1997; Ziegler et al., 2001, and Experiment 3 of Chapter 1) observed a significant effect of length when adult participants encountered the new novel words for the first time, and that the naming reaction time reduced over subsequent blocks (mainly of the long items), with a gradual speeding up of RTs. It is expected to see that the changes in naming RTs would be more apparent in the long items in both orthographic and phonological training conditions. The predictors that affected orthographic and phonological word learning were also explored. Experiment 5 then investigated the role of phonology in orthographic learning in a between-subject design while Experiment 6 minimised the activation of orthographic representations in phonological training by utilizing two types of distractors, namely the orthographic and non-orthographic distractors.

3.2.1 Method

3.2.1.1 Design

Experiment 4 consisted of two parts, the pre-assessment phase and the main experiment.

3.2.1.2 Participants

Forty native speakers of English (20 male, 20 female) aged 18 – 24 (mean age = 19.88, S.D. = 1.28) took part in the experiment. All participants were undergraduate students at the University of York who were either paid with a small payment or received course credit in return. They all had normal or correct-to-normal vision with no history of reading problems.

3.2.1.3 Materials

The materials were identical to those used in Chapter 2 (Experiment 1, section 2.2.1.2), except that based on the result of the pilot testing, four sets (Sets F, G, H and I) of nonwords (96 nonwords) which were matched on accuracy (all above 90 percent) and on initial letters (12 different letters to make the nonwords as different as possible) were chosen to be the experimental items. Reading speed was matched separately for 4- and 7-letter nonwords. The range, mean and standard deviation of the four- and seven- letter experimental items from the pilot study are shown in Table 3.1. All the experimental items of Experiment 4 are shown in Appendix 6. Sixteen additional nonwords (8 four-letter, 8 seven-letter) were chosen for practice trails prior to the main experiment.

3.2.1.4 Auditory stimuli

Four native speakers of British English (2 male, 2 female) who were unknown to participants recorded all the nonwords in Sets F and G (see Appendix 6). All four speakers recorded multiple repetitions of the nonwords. Stimuli were carefully selected in order to optimize the acoustic clarity of nonwords. Speakers were encouraged to read the nonwords in a loud and clear voice. All the stimuli that reached optimum hearing level were then selected with great care to gather the experimental stimulus. All recordings were normalized to have equivalent peak sound energy and voice onset times. The recordings were made in a sound-attenuated booth with a sensitive microphone. Stimuli were digitized with Cool Edit 2000 (www.cooledit.com) and trimmed to length.

Table 3.1. ***Mean and standard deviation of bigram frequency, neighbourhood size and reading speed of the four- and seven-letter nonwords from the pilot study.***

	Nonwords							
	Set F		Set G		Set H		Set I	
	4-letter	7-letter	4-letter	7-letter	4-letter	7-letter	4-letter	7-letter
Log bigram frequency								
Mean	3.19	3.35	3.09	3.29	3.19	3.34	3.23	3.31
S.D.	0.35	0.11	0.30	0.13	0.23	0.15	0.29	0.17
Neighbor-hood size								
Mean	6.08	0.08	6.58	0.08	5.67	0.17	5.25	0.00
S.D.	3.96	0.29	3.90	0.29	3.60	0.58	4.03	0.00
Phonemes								
Mean	3.67	6.00	3.58	5.83	3.75	5.92	3.58	6.08
S.D.	0.49	0.74	0.67	0.58	0.45	0.79	0.51	0.67
Reading speed (RTs in ms) from the pilot study								
Mean	634	745	637	741	636	743	632	754
S.D.	38	77	47	53	55	71	29	63
Range	573 - 695	647 - 899	588 - 733	669 - 811	529 - 738	641 - 836	570 - 688	652 - 856
Naming accuracy in the pilot study								
Mean (max =19)	17.75	17.50	17.00	17.08	17.42	17.42	17.75	17.42
S.D.	0.97	1.67	0.79	1.16	1.16	1.51	1.71	1.08
Range	17-19	14-19	16-18	16-19	15-19	15-19	15-19	16-19

Note. RTs = Reaction time (naming latency) in ms; S.D. = standard deviation

3.2.1.5 Procedure

The experiment began with participants signing a consent form. Participants then completed the 45 minutes pre-assessment phase which was followed by the main experiment.

3.2.1.6 Pre-assessment

All participants completed measures of language ability in one session lasting approximately 45 minutes. Tasks were administered to all participants in the same order.

3.2.1.6.1 *Decoding ability*

Decoding was assessed using the Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) of the Test of Word Reading Efficiency (TOWRE, Torgesen et al., 1999). In this test participants are asked to read a list of words and a list of nonwords of increasing length and difficulty as quickly as they can. Efficiency is indexed by the number of words and nonwords decoded correctly in 45 sec. The test provides norms for individuals aged 6 to 24 years.

3.2.1.6.2 *Vocabulary ability*

Vocabulary was measured using the vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (Wechsler, 1999). This subtest is a measure of expressive vocabulary in which participants are asked to verbally define words. The WASI provides norms for individuals aged 6 to 89 years.

3.2.1.6.3 *Rapid naming*

Rapid naming was measured using the Comprehensive Test of Phonological Processing (CTOPP, Wagner et al., 1999). The Rapid Digit Naming subtest (RDN) task consists of a set of six digits (4, 7, 8, 5, 2, 3) that are displayed in random sequence six times for a total of 72 stimuli. Participants were asked to name the digits from left to right as quickly as possible and the total time to complete the RDN task was recorded.

3.2.1.6.4 *Phonological awareness*

Phonological awareness was measured using the Comprehensive Test of Phonological Processing (CTOPP). One elision subtest of CTOPP was chosen in which a single initial, medial or final phoneme of a word must be deleted and the participant must say what remains. (e.g., deleting the /k/ from "fixed" and responding "fist").

3.2.1.7 Main experiment

Experiment 4 consisted of three parts, the first naming test, the learning phase and the second naming test. The learning phase involved visual training on one set of nonwords (reading aloud) and phonological training on another set (hear-and-repeat). The order of the two forms of training was counterbalanced across participants. Participants began by reading aloud Sets F, G and H (72 nonwords) which were randomly interleaved in the first naming test. Participants then moved on to the learning phase. Training was given on two of the sets (F and G). Half the participants received 8 training blocks in which they read aloud one of the two training sets then 8 training blocks in which they heard and repeated the other training set. The remaining participants received 8 blocks of hear-and-repeat training followed by 8 blocks of reading aloud training. Sets F and G were assigned to visual or phonological training and to first or second training in a counterbalanced way. At the end of training, participants performed the second naming test in which they read aloud interleaved Sets F, G and I. Table 3.2 illustrates the design of the experiment. All of the stimulus of Sets F, G, H, and I are shown in Appendix 6.

Table 3.2. ***Illustration of the distribution of training and test sessions.***

Participants group	**Test 1 (block 1)**	**Training (blocks 2-9)**	**Test 2 (block 10)**
1	read aloud F, G & H	read Set F, repeat Set G	read aloud F, G, & I
2		repeat Set G, read Set F	
3		read Set G, repeat Set F	
4		repeat Set F, read Set G	

3.2.1.7.1 *First naming test (Block 1)*

The participants were tested individually. They were seated approximately 60 cm from the monitor. They wore a set of headphones with a high sensitivity microphone attached. The microphone was linked to a voice key that detected vocal input. Participants were instructed to pronounce the nonwords clearly, and as quickly and accurately as possible, without coughs or hesitations. Before the experimental session, there was a practice session in which 16 items were shown for participants to become familiar with the experimental procedure. It also gave the experimenter the opportunity to adjust the microphone, if necessary.

Nonwords from Sets F, G, and H were presented, and reaction times recorded, using the E-prime software system version 1 (Schneider, Eschman, & Zuccolotto, 2002). On each trial, a black fixation cross appeared in the centre of the screen for 1000 ms. The fixation cross was followed by a nonword displayed in lower case Times New Roman font point 18, which was presented for 2000 ms. Participants were asked to pronounce the nonword as quickly and accurately as possible once they saw the nonword appeared on screen. After the presentation of the nonword, a blank screen was shown for 1000 ms. Subsequently the next trial started. No feedback was given. The test assistant marked the accuracy of each response (right, wrong, or invalid if the voice key had been triggered by another sound).

3.2.1.7.2 *Training (Block 2 – 9)*

After the first naming task (Block 1), half of participants were then trained on the orthography of either Set F or G by reading the stimuli aloud, they were then trained on the phonology of the remaining set by hearing and repeating the stimuli. The remaining participants received 8 blocks of hear-and-repeat training followed by 8 blocks of reading aloud training. All stimuli in the training phase were shown in Powerpoint. No RTs were recorded.

3.2.1.7.3 *Hear-and-repeat training*

Participants were trained on the phonology of either set F or G. On each trial a fixation cross appeared in the centre of the computer screen for 1000 ms, then one of the spoken stimulus items was presented at a comfortable listening level over a professional quality earphone for 5000 ms along with the presentation of ‘XXXX’ or ‘XXXXXXX’ in the centre of the screen depending on the length of the stimuli. Participants then repeated the item they just heard. After the verbal presentation of the nonword, a blank screen was shown for 1000 ms. There were eight blocks in the hear-and-repeat training in which participants heard the 24 nonwords eight times. The order of nonwords was pseudo-randomized across blocks and the order of blocks was fixed between participants. The experimenter sat at the opposite side of the room to monitor the accuracy of nonwords repetition. Fourteen nonwords (7 four-letter, 7 seven-letter) were chosen for practice trials prior to the main experiment.

3.2.1.7.4 *Reading aloud training*

Participants were trained on the orthography of the set by reading the nonword aloud as quickly and accurately as possible once the nonword appeared on screen. The procedure of read aloud training is exactly the same as the first naming task. There were eight blocks in the reading aloud training in which participants read the 24 nonwords for eight times. The order of nonwords was pseudo-randomized across blocks and the order of blocks was fixed between participants. The experimenter sat at the opposite side of the room to monitor the accuracy of nonwords repetition. Fourteen nonwords (7 four-letter, 7 seven-letter) were chosen for practice trials prior to the main experiment.

3.2.1.7.5 *Second naming test (Block 10)*

After being trained for 8 blocks, participants were asked to name Sets F, G and I. The random order procedure was exactly the same as the first naming task.

3.2.2 Result

3.2.2.1 Data trimming

Only RTs for correct responses were analysed. Naming errors, hesitations and failures to activate the voice key were removed from the analysis along with RTs less than 100 ms or longer than 2.5 SDs above the mean (defined separately for each participant in each block and for each length after removal of the very short RTs). Naming errors, hesitations and failures to activate the voice key occurred on 40 trials (0.7% of the total). An additional 60 RTs were removed at the stage of RT trimming (1.0%), leaving 5660 RTs for analysis (98.3% of the total).

Table 3.3 shows the percent and RT results for correct, trimmed responses. Accuracy never fell below 96% correct for any stimulus type in any block of trials. For that reason, the statistical analysis will be confined to the RT data. RT data were analysed by participants. Full details of the statistical analyses are presented in Appendix 6 where effect sizes are reported in terms of the partial eta squared statistic (η_p^2). The main outcomes will be summarised here.

3.2.2.2 Reading accuracy

The non-parametric Wilcoxon test was adopted to compare the effect of Length between groups as accuracy was found to violate the assumption of normality (Kolmogorov-Smirmov test of normality, $p < .05$). Accuracy was generally very high (98.3% correct overall). Ceiling effects meant that there was no significant difference between accuracy for 4- and 7-letter nonwords, $W(40) = 154.00$, $Z = 0.522$, $p = .601$, and no overall difference in accuracy between block 1 and 10, *W*(40) = 203.00, *Z* = 1.597, *p* = .110. Figure 3.5 shows the mean accuracy for each block in both orthographic and phonological training conditions. Table 3.3 shows the percent and RT results for correct, trimmed responses.

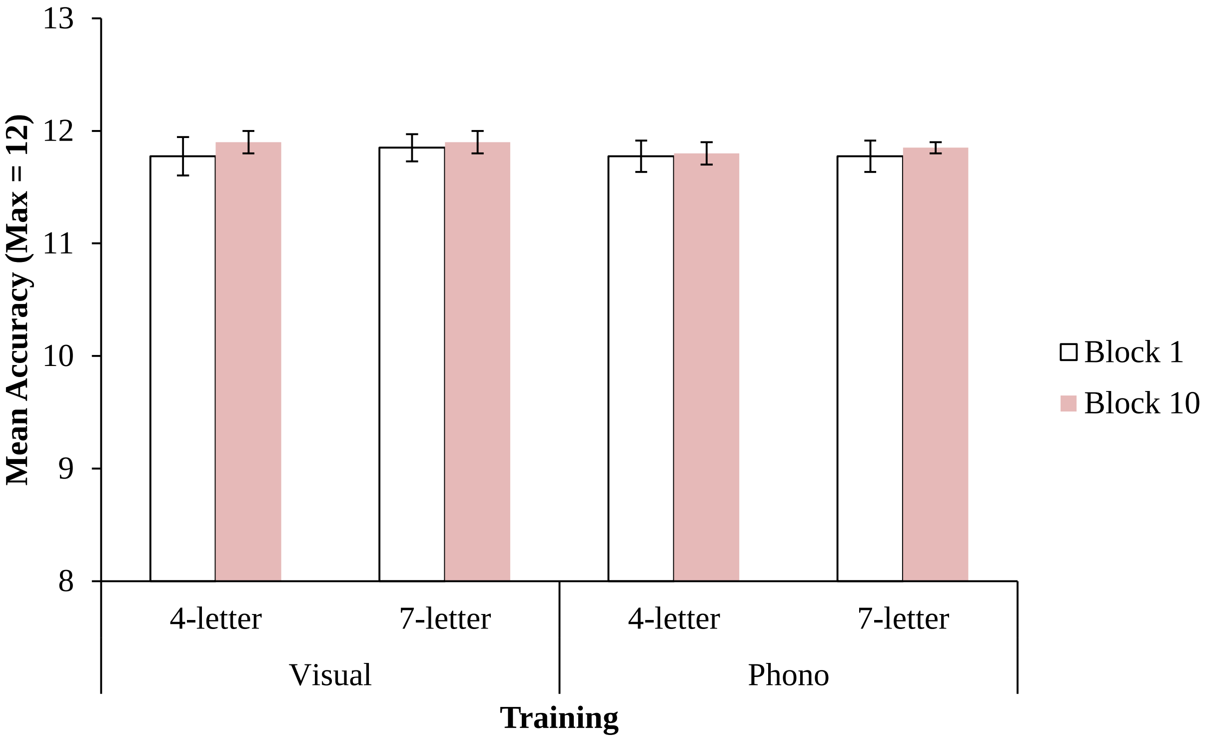

Figure 3.5. **The accuracy of naming 4- and 7-letter nonwords in Blocks 1 and 10 in the visual and phonological training conditions. Error bars show 95% confidence intervals.**

3.2.2.3 Naming latencies (RTs)

Figure 3.6 shows the pattern of RTs for correct, trimmed responses across blocks for visual and phonological trained condition. Inspection of Figure 3.6 indicates that naming latencies were faster in the visual compared to the phonological trained condition. At the start of the experiment, both groups were slower to read 7-letter nonwords aloud than 4-letter. The difference in naming RTs for shorter and longer nonwords reduced with training, but the RTs for shorter and longer items converged more in the visual trained than the phonological trained condition in block 10. The effect of training was also more apparent in the long than short items. Those indications were explored in a series of statistical tests reported in Appendix 2. Separate analysis with Sets and Order were included in ANOVA as a covariate, as none of the main effects and interaction of Sets and Order was significant, the result session will focus on the analysis without Sets and Order.

The fillers Sets H and I were matched to the training Sets F and G based on naming latency and accuracy in the pilot study, however, Sets H and I were not counterbalanced in the experiment. The naming latency and accuracy data for those sets will be stated in this chapter but will not be analysed statistically. The aim of including these sets was to see if naming latency changed simply through practice on the task. The result for Sets H and I will be commented briefly but this question will be addressed in Experiment 5 where untrained sets were used in a fully counterbalanced design.

3.2.2.3.1 *Trained items*

The first set of analyses of the RT data were ANOVAs on trained items conducted by-participants on the data with Groups, Blocks, and Length as factors. There were significant main effects of Blocks (faster overall RTs on block 10 than block 1), and Length (faster overall RTs to 4- than 7-letter nonwords). The significant interaction of Blocks and Group reflected the fact that although the naming RTs were similar in both visual and phonological trained group in Block 1, the naming RTs was faster in the visual trained group in Block 10. The significant interaction of Blocks and Length reflected the length effect was larger in Block 1 than 10. The three-way interaction of Test, Group and Length was marginally significant (p = .082), supporting the trend shown in Figure 3.6 that the visual training exhibited greater improvement in learning compared to the phonological training condition. The results were explored further by means of separate analyses of ANOVA and t-test of 4- and 7-letter nonwords in Block 1 and 10.

Table 3.3. ***Mean latencies of correct, trimmed responses, standard deviation (S.D.), and percent correct responses for 4- and 7-letter nonwords in Session 1 and 2 in Experiment 4.***

Test		**1**	**2**
Blocks		**1**	**10**
Visual training			
4-letter nonwords			
	Mean RT	530	506
	S.D.	101.6	93.0
	% correct	100.0	99.1
7-letter nonwords			
	Mean RT	640	546
	S.D.	143.6	112.9
	% correct	97.9	99.1
Phonological training			
4-letter nonwords			
	Mean RT	529	518
	S.D.	102.7	99.6
	% correct	99.0	98.3
7-letter nonwords			
	Mean RT	624	566
	S.D.	142.4	113.1
	% correct	97.9	98.7
Set H			
4-letter nonwords			
	Mean RT	535	
	S.D.	104.2	
	% correct	98.3	
7-letter nonwords			
	Mean RT	637	
	S.D.	141.5	
	% correct	98.8	
Set I			
4-letter nonwords			
	Mean RT		515
	S.D.		99.3
	% correct		98.1
7-letter nonwords			
	Mean RT		618
	S.D.		141.4
	% correct		95.9

Note. RT = Reaction time (naming latency) in ms; S.D. = standard deviation

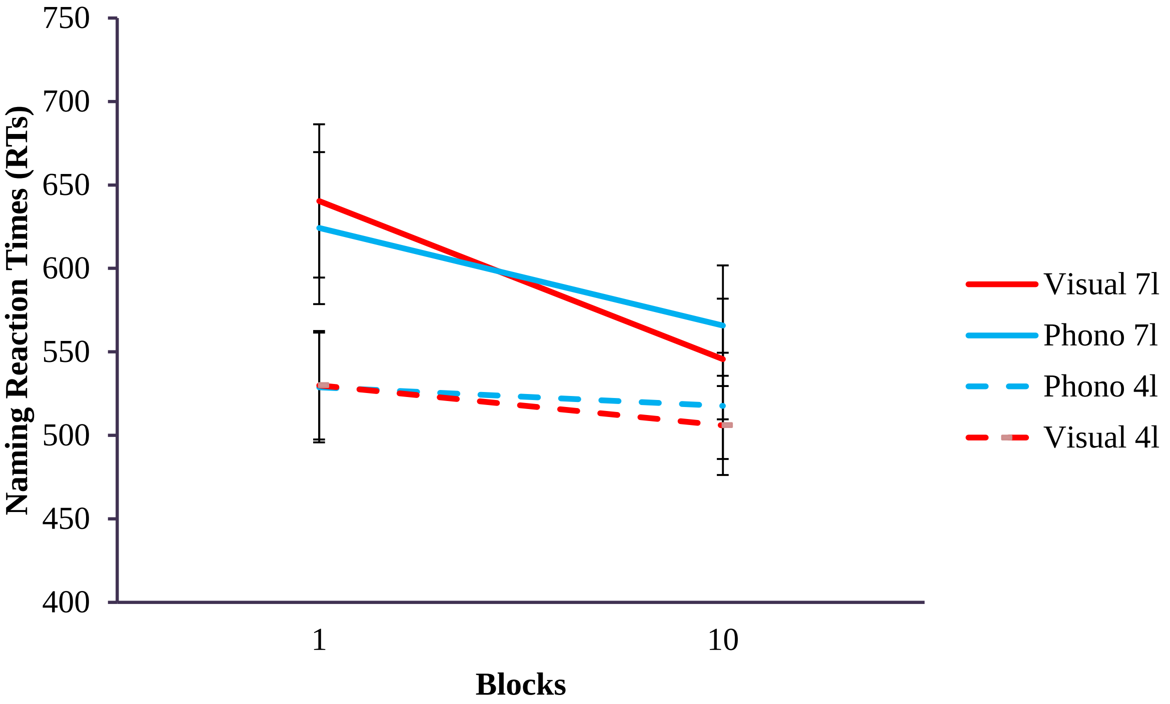

***Figure 3.6.* The naming reaction time (RTs) for 4- and 7-letter nonwords in Blocks 1 and 10 in visual and phonological training conditions of Experiment 4. Error bars show 95% confidence intervals.**

3.2.2.3.2 *4-letter*

The RT data for 4-letter was analysed with Blocks and Groups as factors. The main effect of Blocks was not significant. The interaction of Blocks and Groups was only marginally significant. The interactions were analysed further by means of separate t-test analyses of RTs for 4-letter items in Block 1 and 10. Bonferroni-corrected t-tests ($\propto$ = .01) found no significant difference for visual and phonological trained 4-letter nonwords naming RTs in blocks 1 and 10.

3.2.2.3.3 *7-letter*

The next analysis focused on performance in 7-letter nonwords with factors once again of Blocks and Groups. The main effect of Blocks (faster naming RTs in Block 10 than 1) was significant. The interaction of Blocks and Groups was also significant, reflecting the visual trained condition showed a bigger improvement compared to the phonological trained condition. The interactions were analysed further by means of separate t-test analyses of RTs for 7-letter items in Block 1 and 10. Bonferroni-corrected t-tests ($\propto$ = .01) found a significant difference for 7-letter items in both visual and phonological trained condition in Block 1 and 10. Bonferroni-corrected t-tests ($\propto$ = .01) also found a significant difference of visual and phonological trained 7-letter items in Block 10.

3.2.3 Predictors of initial nonword reading speed and novel word learning

The final set of analyses brought together performance on the test battery with two aspect of their naming latency data: a) RTs to 7-letter nonwords in block 1 of day 1 as a measure of nonlexical reading skill and b) the change in RTs to 7-letter nonwords from block 1 to block 10 on day 1 as a measure of novel word learning in visual and phonological training.

The number of predictor variables was reduced before the regression analyses were run, and some of the variables were transformed to improve the normality of their distributions. There were high correlations among the two word and nonword reading test (*r*s = .54, *p* < .01). A Literacy composite score was calculated for each participant by standardizing and summing the sub-test scores from the TOWRE Sight Word Efficiency (SWE), and Phonemic Decoding Efficiency (PDE).

Univariate normality was tested for each predictor and the dependent variables (RTs to 7-letter nonwords in blocks 1 and 10 of day 1). Sight Word Efficiency, Phonemic Decoding Efficiency, Phonological awareness, and Rapid Digit Naming were found to violate the assumption of normality (Kolmogorov-Smirmov test of normality, *p* < .05). Distributions approximated normality most closely when the Literacy composite score and Phonological awareness were reverse then square root transformed. Sight word efficiency and Rapid digit naming were log transformed. Phonemic Decoding Efficiency was square root transformed. RTs were log transformed to reduce skew.

Table 3.4 shows the correlations among the final predictor variables; also between the predictor variables and RTs to 7-letter nonwords in block 1 of day 1. There were significant correlations among all the predictor variables. All of the predictors except vocabulary correlated significantly with RT, with Phonemic Decoding Efficiency showing the highest correlation, followed by Literacy, Phonological awareness, Sight word efficiency and Rapid digit naming. Multicollinearity among the predictor variables was assessed using the variance inflation factor (VIF). VIF scores of less than 4 indicate that the result will not significantly influence the stability of the parameter estimates (Myers, 1990; Olague, Etzkorn, Gholston, Quattlebaum, 2007). VIF scores for the predictor variables ranged between 1.03 and 1.72.

Table 3.4. ***Correlations among the predictor variables, and between the predictor variables and naming RTs for 7-letter nonwords in block 1 of day 1.***

Variable	1 Vocab	2 Literacy	3 Phon	4 RDN	5 RT
1. Vocabulary	–				
2. Literacy composite	-.138	–			
3. Phonological Awareness	-.422**	.427**	–		
4. Rapid Digit Naming	.018	-.583**	-.267	–	
5. Block 1, 7-letter RTs	-.259	.583**	.531**	-.365*	–

Note. *$p < .05$, **$p < .01$. Note that Phonological awareness and Literacy composite score was reverse then square root transformed. Rapid Digit Naming and Sight Word Efficiency were log transformed. Phonemic Decoding Efficiency was square root transformed. RTs were log transformed to reduce skew.

Linear mixed effects modelling was used to explore the ability of Vocabulary, Literacy, Phonological awareness, and Rapid digit naming to predict initial nonword reading speed and novel word learning. Linear mixed effects (LME) methods analyse all the available data and do not rely on averaging across participants or across items. It allows differences in the baseline performance among participants and items (*random effects*) to be separated from the effects of the predictor variables (*fixed effects*) (Baayen, 2008; Bates et al, 2007; Jones et al., 2008). The analyses were conducted in R using the lme4 (Bates, Maechler, & Bolker, 2012) and languageR (Baayen, 2009) packages.

3.2.3.1 Predicting initial nonword reading speed

The contribution of each predictor variable to predicting RTs for 7-letter nonwords presented in block 1 of day 1 was evaluated by using likelihood ratio tests to compare models that contained all the fixed and random effects with a sequence of models in which the different predictor variables were removed one at a time. These analyses showed that Literacy, $\chi^2(8) = 7.64$, $p < .01$; $\beta = 0.02$, $t = 2.90$, $p < .01$, and Phonological awareness, $\chi^2(8) = 3.96$, $p < .05$; $\beta = 0.06$, $t = 2.04$, $p < .05$, made a significant independent contribution to predicting nonword naming speed. In contrast, Vocabulary, $\chi^2(8) = 0.29$, $p = .592$, and Rapid digit naming, $\chi^2(8) = 1.00$, $p = .999$, made no independent contributions. A similar pattern of results (prediction of initial naming RTs by Phonemic decoding efficiency, $\chi^2(9) = 10.87$, $p < .01$; $\beta = -0.12$, $t = -3.54$, $p < .01$, and Phonological awareness, $\chi^2(9) = 3.60$, $p = .058$; $\beta = 0.05$, $t = 1.94$, $p = .059$), but not Sight word efficiency, $\chi^2(9) = 3.41$, $p = .122$, was obtained when the data was analysed with Literacy replaced by Sight word efficiency and Phonemic decoding efficiency.

3.2.3.2 Predicting learning

Visual and phonological word learning was assessed in terms of the change in naming RTs for 7-letter nonwords between blocks 1 and 10 of day 1. RTs from both blocks were entered into the analysis separately for visual and phonological training. A categorical variable of Time was created to reflect the change in RTs between blocks 1 and 10. A set of predictor variables were then created which were the interactions between Time and Vocabulary, Literacy, Phonological awareness and Rapid digit naming. This makes it possible to evaluate the contribution of each independent variable to predict change in naming RTs to the 7-letter nonwords across blocks independently for visual and phonological training (Field, 2012; Shek & Ma, 2011). A categorical variable of Order was also included in order to take into account that half of the participants have visual training before receiving phonological training.

3.2.3.3 Visual training

The effect of the categorical variable of Time was significant, $\chi^2(10) = 104.76$, $p < .001$, reflecting the reduction in RTs from block 1 to block 10. The effects of the interactions of Time with Vocabulary, $\chi^2(14) = 12.32$, $p < .001$; $\beta = 0.004$, $t = 3.52$, $p < .001$, and Time with Rapid Digit Naming, $\chi^2(14) = 4.17$, $p < .05$; $\beta = 0.16$, $t = 2.05$, $p < .05$, were significant. The interactions of Time with Literacy, $\chi^2(14) = 0.01$, $p = .941$, and Phonological awareness, $\chi^2(14) = 0.49$, $p = .484$, made no independent contributions to predicting RTs change across blocks in visual training. A similar pattern of results (prediction of RTs change by Vocabulary, $\chi^2(16) = 11.29$, $p < .001$; $\beta = 0.004$, $t = 3.37$, $p < .001$, and Rapid digit naming, $\chi^2(16) = 4.52$, $p < .05$; $\beta = 0.18$, $t = 2.13$, $p < .05$) was obtained when the data was analysed with Literacy replaced by Sight word efficiency and Phonemic decoding efficiency.

3.2.3.4 Phonological training

The effect of the categorical variable of Time was significant, $\chi^2(10) = 39.05$, $p < .001$, reflecting the reduction in RTs from block 1 to block 10. The effects of the interactions of Time with Vocabulary, $\chi^2(14) = 23.33$, $p < .001$; $\beta = 0.005$, $t = 4.86$, $p < .001$, was significant. The interactions of Time with Rapid digit naming, $\chi^2(14) = 0.09$, $p = .759$, Literacy, $\chi^2(14) = 0.23$, $p = .634$ and Phonological awareness, $\chi^2(14) = 0.39$, $p = .533$, made no independent contributions to predicting RTs change across blocks in phonological training. A similar pattern of results (prediction of RTs change by Vocabulary, $\chi^2(16) = 19.14$, $p < .001$; $\beta = 0.005$, $t = 4.40$, $p < .001$) was obtained when the data were analysed with Literacy replaced by Sight word efficiency and Phonemic decoding efficiency.

3.2.4 Discussion

Similar to the results of Chapter 2, pre-selection of the items for Experiment 4 on the basis of the pilot study meant that accuracy of reading the nonwords was high in both orthographic and phonological training conditions. Ceiling effects meant that there was no detectable influence of length or blocks on accuracy.

Though a different set of nonwords was used in this experiment compared to those in Chapter 2, the pattern of result was quite similar to that in Experiment 3 of Chapter 2. Naming latencies to nonwords seen for the very first time in block 1 were 530 ms for 4-letter nonwords and 632 ms for 7-letter nonwords (the average mean of both training conditions). This is in line with the literature of the length effect in nonwords naming in English (Weekes, 1997; Ziegler et al., 2001).

In both orthographic and phonological training conditions, RTs became shorter across blocks as the nonwords became familiar. This result was more apparent for the 7-letter nonwords. Mean RTs for 4-letter nonwords reduced by a non-significant 24 ms across the 10 blocks of orthographic training while the mean RTs for 7-letter nonwords reduced by 94 ms. Mean RTs for 4-letter nonwords reduced by 11 ms across the 10 blocks of phonological training while the mean RTs for 7-letter nonwords reduced by 58 ms. The convergence of RTs to short and long nonwords is shown in *both* training conditions in Figure 3.6. In accordance to the ‘lexical quality hypothesis’, learning was better in the orthographic training conditions which participants had the benefit of receiving training on both the orthography and the phonology of the stimuli. Yet, the fact that the convergence of RTs to shorter and longer nonwords was also shown in the phonological training condition implies that training the phonology of the new words is sufficient to build representations in adults’ mental lexicons. This is in line with previous developmental (Reitsma, 1983) and adult literature (Chalmers and Burt, 2008; Sandak et al., 2004) that as phonology is an essential part of learning new words, training the phonology of new words is sufficient to help participants to build representations in the mental lexicon.

Linear mixed effects modelling found that Literacy (composite score of TOWRE SWE and PDE) and phonological awareness made a significant contribution to predicting 7-letter nonwords naming speed in block 1 even when vocabulary and rapid digit naming were taken into account. PDE and phonological awareness were still the significant predictors of nonword reading speed even when Literacy was replaced by SWE and PDE. This result is consistent with the developmental (e.g. Bowyer-Crane et al, 2008; Muter et al., 2004; Ricketts et al., 2009) and adult (Young et al., 2002) studies showing that phonological skills is still a crucial factor that affect the speed of reading nonwords when children proceeds to their adulthood. This finding is also in line with Chalmers and Burt (2008), Sandak et al. (2004) and Pennington et al. (1987) that it challenges the ‘phonological bypass hypothesis’ as it illustrates that phonological encoding skills still play an essential role in nonword reading in adults.

As reported by Braze et al. (2007), and Nation and Snowling (2004), expressive vocabulary is a significant predictor of the change in naming RTs for 7-letter nonwords between block 1 and 10 of day 1 in both orthographic and phonological training conditions even when literacy and phonological awareness were taken into account. The result extends Ouellette’s (2006) finding that vocabulary depth/expressive vocabulary is not only related to reading comprehension but also affect how well participants can build representations in the mental lexicon.

Savage et al.’s (2005) regression analysis revealed that the significant predictor of reading rate, which is based on the number of words read per minute, was digit naming speed rather than picture naming speed. Even after further controlling reading accuracy, digit naming was a significant predictor of reading rate whereas phonological awareness tasks predicted reading accuracy and comprehension. In Experiment 4, rapid digit naming only contributed to the variance in the change of naming RTs for 7-letter nonwords between block 1 and 10 of day 1 in the orthographic, but not the phonological training condition. Consistent with Pennington and Lefly (2001) and Young and Bowers (1995), the current study also demonstrated that RAN and phonological awareness predict different aspects of reading ability. As mentioned in section 1.4.2 of Chapter 1, the significant contribution of RAN to visual word learning may be due to the visual stimuli in the task (in this experiment it is the digits, e.g. 8) have to be mapped rapidly to their names (i.e., eight). This process is particularly similar to the procedure in the orthographic training condition in this experiment which participants have to map the orthography of a nonword to its phonology. This may imply that participants with better rapid digit naming skills have a better lexical access to both the orthographic input lexicon and phonological output lexicon which may contribute to the efficient orthographic learning. The mechanism of how phonological skill, vocabulary and RAN ability supports word learning will be further discussed in the General Discussion (section 3.5).

3.3 Experiment 5: The role of phonology in orthographic learning (between subject design)

Previous studies have found the training that participants received in task 1 may contribute to the cross-task correlation that may affect their performance in task 2 (Lovett, Daily, & Reder, 2000). This implies that half of the participants in Experiment 4 may tend to activate orthographic codes in the phonological training condition after being trained in the orthographic training conditions for 9 blocks. In order to control for this cross-task correlation, Experiment 5 adopted a between-subject design in which participants are trained either in the orthographic or in the phonological learning condition.

Experiment 5 also tries to demonstrate that the learning effect that was observed in Experiment 4 was not due to a specific set. Three sets of nonwords (G, H, and I; same items in the equivalent sets in Experiment 4) were adopted, with each set containing 12 4-letter and 12 7-letter items (as in Experiment 4). Each participant received one set of nonwords in all 10 blocks of the experiment, with a second set shown in block 1 only and a third set shown in block 10 only. As the order of sets was counterbalanced across participants, the question of interest is whether the learning effect in both the orthographic and phonological training condition will be greater than the general improvement that is obtained from the untrained sets.

3.3.1 Method

3.3.1.1 Participants

Forty-eight native speakers of English (24 male, 24 female) aged 18 – 26 (mean age = 19.46, S.D. = 1.68) took part in the experiment. All participants were undergraduate students at University of York who were either paid with a small payment or received course credit in return. They all had normal or correct-to-normal vision with no history of reading problems.

3.3.1.2 Materials

The materials were identical to those used in Experiment 4, except that only Sets G, H and I were used in this experiment. All the experimental items for Experiment 5 are shown in Appendix 6.

3.3.1.2.1 *Auditory stimuli*

Two native speakers (different from Experiment 4) of British English (1 male, 1 female) who were unknown to participants recorded all the nonwords in Sets G, H, and I (see Appendix 6). Other settings of the auditory stimuli were identical to Experiment 4.

3.3.1.3 Procedure

Both reaction time (RTs) and accuracy were measured for the experiment. The session began with participants signing a consent form. The main experiment consisted of three parts, the first naming test, the learning phase and the second naming test. The learning phase involved participants either having visual training on one set of nonwords (reading aloud) or phonological training on another set (hear and repeat). Participants began by reading aloud Sets G and H (48 nonwords, with set G serving as the fillers for the first naming test.) which were randomly interleaved in the first naming test. Participants then moved on to the learning phase. Training was given on set H. Participants in the visual training group received 8 training blocks in which they read aloud one of the three training sets. Participants in the phonological training group received 8 training blocks in which they received hear-and-repeat training. At the end of training, participants performed the second naming test in which they read aloud interleaved Sets H, and I, with set I serving as the fillers for the second naming test. Table 3.5 illustrates the design of the group that received visual training in the experiment. Table 3.6 illustrates the design of the group that received phonological training in the experiment. The order of Sets G, H, and I were fully counterbalanced.

Table 3.5. ***Illustration of the distribution of training and test sessions in the group that received visual training.***

Group	Block 1 (Sets)	Block 2 – 9 (Set)	Block 10 (Sets)
1	Read aloud G + H	Read aloud H	Read aloud H + I
2	Read aloud G + I	Read aloud I	Read aloud I + H
3	Read aloud H + G	Read aloud G	Read aloud G + I

Table 3.6. ***Illustration of the distribution of training and test sessions in the group that received phonological training.***

Group	Block 1 (Sets)	Block 2 – 9 (Set)	Block 10 (Sets)
1	Read aloud G + H	Hear and repeat H	Read aloud H + I
2	Read aloud G + I	Hear and repeat I	Read aloud I + H
3	Read aloud H + G	Hear and repeat G	Read aloud G + I

3.3.2 Result

3.3.2.1 Data trimming

Only RTs for correct responses were analysed. Naming errors, hesitations and failures to activate the voice key were removed from the analysis along with RTs less than 100 ms or longer than 2.5 SDs above the mean (defined separately for each participant in each block and for each length after removal of the very short RTs). Naming errors, hesitations and failures to activate the voice key occurred on 41 trials (0.9% of the total). An additional 65 RTs were removed at the stage of RT trimming (1.4%), leaving 4502 RTs for analysis (97.7% of the total). Table 3.7 shows full accuracy and RT results for correct, trimmed responses. Accuracy never fell below 95% correct for any stimulus type in any block of trials. For that reason, the result section will mainly focus on the statistical analysis of the RT data. RT data were analysed by participants. Full details of the statistical analyses are presented in Appendix 2 where effect sizes are reported in terms of the partial eta squared statistic (η_p^2). The main outcomes will be summarized in this result section.

3.3.2.2 Reading accuracy

Non-parametric Mann-Whitney *U* test was adapted to compare the effect of Length between groups as accuracy was found to violate the assumption of normality (Kolmogorov-Smirmov test of normality, $p < .05$). Accuracy was generally very high (overall mean 97.7% correct and never below 95.1%). Ceiling effects meant that there was no significant difference between accuracy for 4-letter nonwords, $U(48) = 312.00$, $Z = .589$, $p = .556$; 7-letter nonwords, $U(46) = 265.00$, $Z = -.537$, $p = .591$; Block 1, $U(46) = 302.00$, $Z = .294$, $p = .769$; and Block 10, $U(46) = 302.00$, $Z = .294$, $p = .769$, among the two training conditions. Figure 3.7 shows the mean accuracy for the trained and untrained items for each block in both orthographic and phonological training conditions.

3.3.2.3 Naming latencies (RTs)

Figure 3.8 shows the pattern of RTs for correct, trimmed responses to trained and untrained items across blocks for orthographic and phonological learning conditions. Inspection of Figure 3.8 indicates that naming latencies were faster in the visual compared to the phonological learning condition. At the start of the experiment, both groups were slower to read 7-letter nonwords aloud than 4-letter. The difference in naming RTs for shorter and longer nonwords reduced with training, but the RTs for shorter and longer items converged more in the visual than the phonological learning condition in block 10. The effect of training was also more apparent in the long than short items. Those indications were explored in a series of statistical tests reported in the Appendix 2. The analysis of the RT data was done in two parts – first a global analysis of RTs in trained and untrained nonwords across blocks 1 to 10 and second a separate analysis of RTs to visual and phonological learning conditions.

Table 3.7. ***Mean latencies of correct, trimmed responses, standard deviation (S.D.), and percent correct responses for 4- and 7-letter nonwords in block 1 and 10 in Experiment 5.***

Blocks		**1**	**10**
Visual training			
4-letter nonwords			
	Mean RT	548	494
	S.D.	104.1	77.4
	% correct	97.9	98.3
7-letter nonwords			
	Mean RT	608	504
	S.D.	124.7	92.0
	% correct	97.6	97.9
Visual Fillers			
4-letter nonwords			
	Mean RT	541	492
	S.D.	96.6	90.5
	% correct	97.9	97.6
7-letter nonwords			
	Mean RT	620	570
	S.D.	143.8	119.5
	% correct	97.9	96.9
Phonological training			
4-letter nonwords			
	Mean RT	538	540
	S.D.	97.4	73.6
	% correct	98.3	99.0
7-letter nonwords			
	Mean RT	611	568
	S.D.	130.4	81.2
	% correct	96.5	96.9
Phonological fillers			
4-letter nonwords			
	Mean RT	550	548
	S.D.	87.1	72.5
	% correct	99.1	98.3
7-letter nonwords			
	Mean RT	607	632
	S.D.	116.0	109.8
	% correct	98.3	95.1

Note. RT = reaction time (naming latency); S.D. = standard deviation

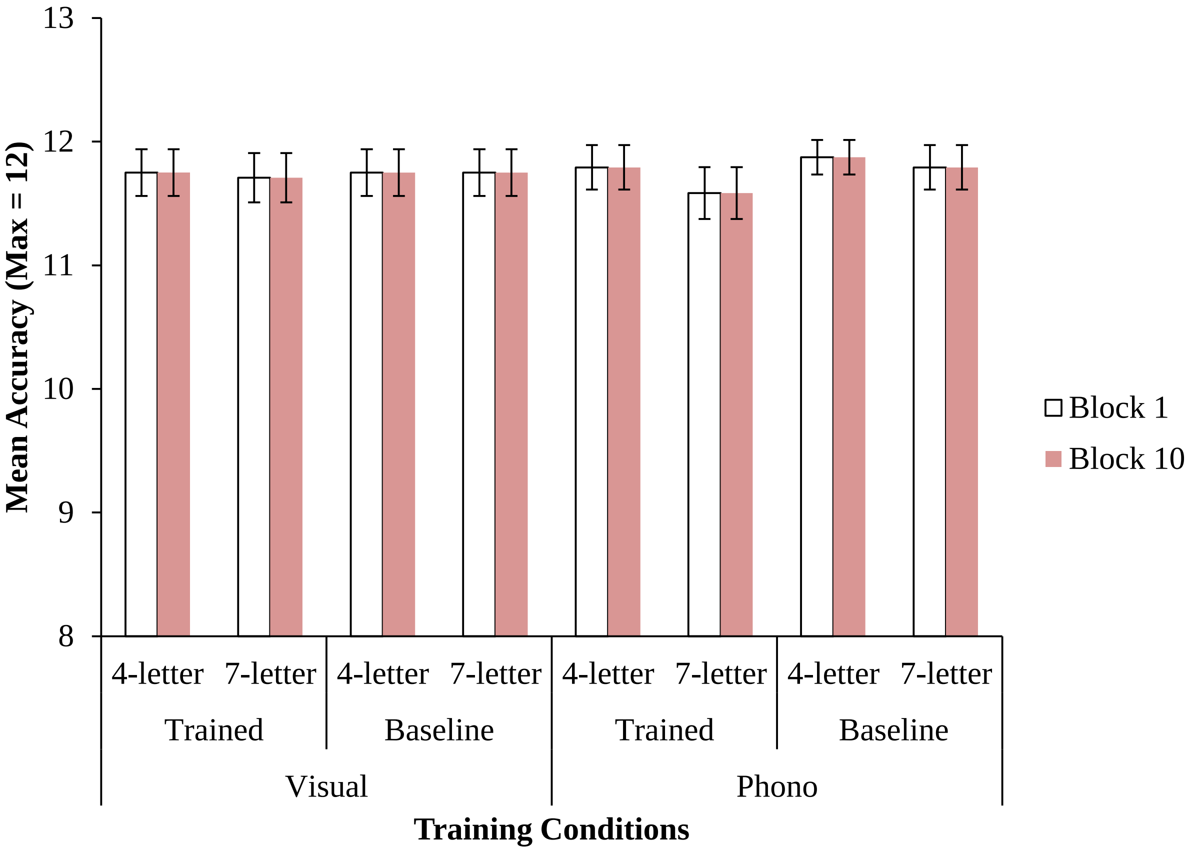

***Figure 3.7.* The accuracy of naming 4- and 7-letter nonwords in Blocks 1 and 10 in the visual and phonological Trained and Baseline conditions.**

3.3.2.3.1 *Global analysis*

The first set of analyses of the RT data were global ANOVAs conducted by-participants on the data for both sessions with Group (visual vs phonological training), Training (train vs filler items), Blocks (1 vs 10) and Length (4 vs 7) as factors. There were significant main effects of Training (faster RTs for the trained than the fillers items), Length (faster overall RTs to 4- than 7-letter nonwords), and Blocks (RTs decreasing across blocks). The majority of the interactions were significant, including the interactions involving Group, supporting the indications in Figure 3.8 that the pattern of results for naming latencies was different in visual and phonological learning conditions. The results were explored further by means of separate analyses of visual and phonological learning conditions. The significant interaction of Group and Blocks reflected the fact that naming RTs were faster in visual learning group in Block 10 than Block 1. The significant interaction of Training and Length illustrated the length effect was smaller in the trained than filler items. The significant interaction of Length and Blocks showed the length effect was smaller in Block 10 than 1. The significant interaction of Training, Length and Blocks illustrated the effect of length was smaller in the trained items in Block 10 than 1. The significant interaction of Group, Training and Length reflected the length effect was smaller in the trained condition of the visual learning condition. This interaction was further investigated with separate analysis of visual and phonological learning conditions.

3.3.2.3.2 *Visual learning group*

The RT data for the visual learning group was analysed with Training, Blocks and Length as factors. There were significant main effects of Training (slower overall RTs in the fillers than the trained items), Blocks (decrease in RTs across blocks) and Length (faster RTs to 4- than 7-letter items). All of the interactions were significant. The significant interaction of Training and Length showed that there was a smaller length effect in the trained items than the filler items. The significant interaction of Training and Blocks reflected the naming RTs of the trained items was faster in Block 10 than 1. The significant interaction of Length and Blocks showed there was a smaller length effect in Block 10 than Block 1. The three way interaction between Training Blocks and Length indicates that though there was an apparent length effect in both trained and filler items in Block 1, after receiving read-aloud training for eight blocks, there was a greater reduction in the effect of length for the trained than the untrained nonwords in block 10.

To further explore the aforementioned interactions, RTs in blocks 1 and 10 were analysed separately with factors of Training (trained vs. untrained) and Length. In block 1, the main effect of Length was significant but the main effect of Training and the Training x Length interaction were not significant (as none of the items has undergone any training so effects of ‘Training’ would not be expected.) By block 10 the trained items have been seen in each of the 9 previous blocks the untrained items are new. In block 10 the main effects of Training (Naming RTs of trained items faster than untrained items) and Length were both significant. The training x Length interaction was also significant, reflecting the fact that the effect of length in block 10 was 78 ms for untrained items but only 10ms for trained items, and the difference between trained and untrained items was 2 ms for 4-letter nonwords compared with 66 ms for 7-letter nonwords. In Bonferroni-corrected *t*-test ($\propto$ = .0125) the difference between the naming RTs between block 1 and 10 was significant for trained 4-letter items, trained 7-letter items, untrained 4-letter items and untrained 7-letter items. This replicates the finding of Experiment 2 in Chapter 2 that though there was some general improvement for reading filler items in Block 10 after participants had been trained for reading aloud for 9 blocks, the substantial reduction of the length effect could only be observed with the trained items. (see section 2.3.2 in Chapter 2).

3.3.2.3.3 *Phonological learning group*

As in the visual learning group, the majority of main effects and interactions were significant. The main effect of Training (slower overall RTs in the fillers than the trained items), and Blocks (decrease in RTs across blocks) were significant. The significant interaction of Training and Length illustrated there was a smaller effect of length in the trained compared to filler items. The significant interaction between Training and Blocks reflected that though the naming RTs of both trained and filler items was similar in Block 1, the RTs of trained items was faster in Block 10. The significant three-way interaction of Training, Length and Blocks reflected the fact that RTs decreased between blocks 1 and 10 with a greater reduction in the effect of length for the trained than the untrained nonwords.

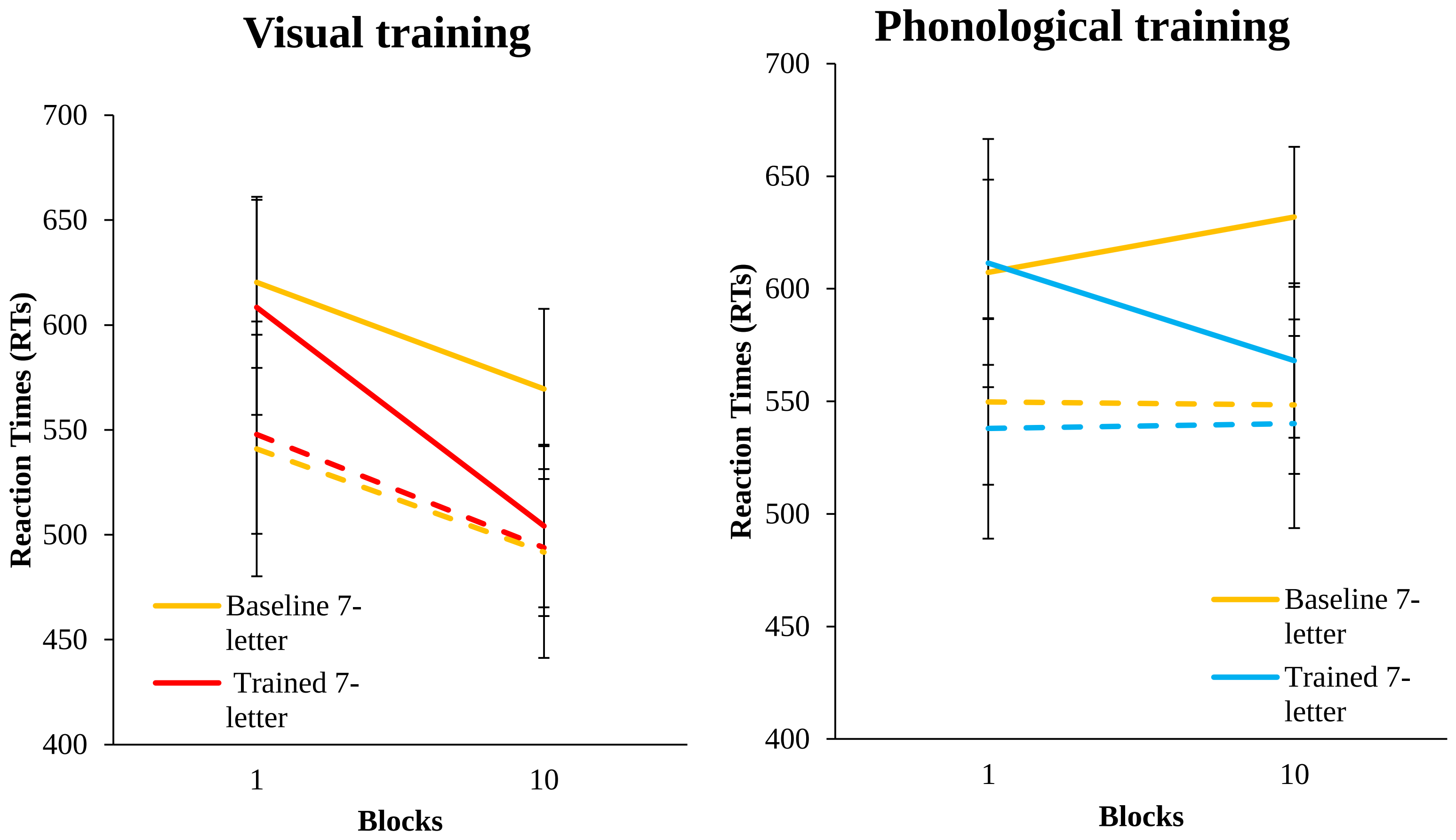

Figure 3.8. **The naming reaction time (RTs) for 4- and 7-letter nonwords in Blocks 1 and 10 in visual and phonological training conditions of Experiment 5. Error bars show 95% confidence intervals.**

To further explore these interactions, RTs in blocks 1 and 10 were analysed separately with factors of Training and Length. In block 1, the main effect of Length was significant but the main effect of Training and the Training x Length interaction were not significant. In block 10 the main effects of Training and Length were both significant. The Training x Length interaction was also significant, reflecting the fact that the effect of length in block 10 was 84 ms for untrained items but only 28 ms for trained items, and that the difference between trained and untrained items was 8 ms for 4-letter nonwords compared with 64 ms for 7-letter nonwords. In Bonferroni-corrected *t*-test ($\propto = .0125$) the difference between the naming RTs between block 1 and 10 was marginally significant for trained 7-letter items ($p = .087$), but not for trained 4-letter, untrained 4-letter and untrained 7-letter.

3.3.3 Discussion

RTs to trained nonwords in both the orthographic and phonological training conditions showed a similar pattern to that seen in Experiment 4, with both conditions showing a larger reduction of length effect for the trained compared to the fillers items. Training on the phonology of novel words had successfully built representations in the mental lexicon, yet, orthographic training had yielded greater improvement compared to the phonological training condition. This is evidenced by the fact that the three-way interaction of Training, Length and Blocks was significant in *both* orthographic and phonological training conditions. The result of this experiment confirmed that learning the phonology of the novel word is still an essential part of learning new words in adults (Chalmers and Burt, 2008; Sandak et al., 2004). This issue will be further elaborated in the General Discussion (section 3.5).

However, there are differences between the two training conditions that were evidenced by the interactions involving Group in the global analysis. While all the post-hoc comparison of RTs between Block 1 and 10 in the orthographic training condition were significant for 1) trained 4-letter, 2) trained 7-letter, 3) untrained 4-letter and 4) untrained 7-letter (all reached $p < .01$), only the trained 7-letter t-test was marginally significant ($p = .087$) in the phonological training condition. No significant improvement was shown for both 4- and 7-letter fillers items in the phonological training condition. This demonstrates that participants showed a greater improvement in the novel words (for both trained and filler items) in the orthographic training condition.

The effect of training was also more apparent in the long compared to the short items. Of note is the fact that the RTs pattern in the orthographic training condition echoes with the result in Experiment 2 of Chapter 2. Though the RTs to the fillers items also fell between block 1 and block 10, that reduction was not, however, as great as for the trained nonwords and the effect of length in block 10 remained at 78 ms for the fillers items compared with 10 ms for the trained items. Given that the improvement in the fillers items was only noticeable in the orthographic training condition, but not the phonological training condition; this may imply that the blocking/list context effect that was mentioned in Section 2.3.3. of Chapter 2 was specific to visual training but not phonological training.

3.4 Experiment 6: Phonological learning with distractors

Experiment 5 showed that training the phonology of a new novel word is sufficient to build lexical entries in the mental lexicon that can help them to distinguish trained and non-trained words in the reading aloud task after training. Yet, Bürki et al. (2012) found orthographic activation of French novel words by using a phonological learning task. Participants learned the auditory forms of potential reduced variants of novel French words (e.g. /pluR/) and their semantic meaning with pictures of novel objects over 4 days. After the fourth day of training, the spelling of each novel word was presented once. Half the words were spelled with an orthographic representation of the schwa (i.e., ‘e’), half were not. They then examined whether production latencies to reduced variants whose spelling contained an orthographic representation of the schwa were longer than production latencies to the same novel words with no representation of schwa in the spelling. The longer latencies observed for novel words with an internally attested cluster and an ‘e’ in the spelling suggest that participants had stored these novel words with a schwa and a non-schwa representation when their spellings contained an orthographic representation of the schwa, and that these two representations compete during a reading aloud task.

Given that some studies (e.g. Bürki et al., 2012) findings suggest that the influence of spelling occurs because the orthographic information is automatically and mandatorily activated on-line whenever listeners process a spoken word/novel word and this effect happened rapidly in both direction (i.e., orthography to phonology), Kyte and Johnson (2006) tried to compress phonological recoding while allowing orthographic processing to occur by asking participants to say 'LA' from the onset of presentation of novel words. Maloney et al. (2009) tried to minimize the phonological processing in orthographic learning by using the case decision task in which participants were asked to verbally identify the case in which a letter string was presented by responding 'upper' or 'lower' aloud.

This experiment aims to attenuate orthographic activation during phonological training by incorporating orthographic (letter strings) and non-orthographic (pictures) distractors in the hear-and-repeat condition. There were four conditions in this experiment, where the read aloud and hear-and-repeat conditions were the same as those in Experiment 4 and 5, Experiment 6 included two additional conditions, namely 1) Hear-and-repeat with orthographic distractors in which the 4-letter strings would change every 500 ms on the screen and 4) Hear-and-repeat with non-orthographic distractors in which facial pictures would change every 500ms on the screen. In order to ensure that participants look at the distractors (rather than ignoring them), a red dot appeared on the screen in 6% of the experimental session. Participants pressed a button of a response box when it appeared. The first block of reading aloud training in Experiment 4 and 5 had also been eliminated in Experiment 6 in order to reduce the possibility that participants use the spelling of the auditory novel words as a strategy to learn them. As most of the training effect was observed in the long items rather than the short items in Experiment 4 and 5, only long items (7-letter nonwords) are used in Experiment 6.

3.4.1 Method

3.4.1.1 Design

This study is run as a between-subject design.

3.4.1.2 Participants

One hundred and four native speakers of English (48 male, 56 female) aged 18 – 24 (mean age = 19.41, S.D. = 1.31) took part in the experiment. All participants were undergraduate students at the University of York who were either paid with a small payment or received course credit in return. They all had normal or correct-to-normal vision with no history of reading problems.

3.4.1.3 Materials

Based on the result of the previous experiment, two new sets of nonwords (Sets J and K, 30 nonwords) which were matched on accuracy (all above 92 percent) and on initial letters (12 different letters to make the nonwords as different as possible) were chosen from Set F, G, H and I to be the experimental items. The range, mean and standard deviation of the seven- letter experimental items were shown in Table 3.8. All the experimental items of Experiment 6 are shown in Appendix 6. Twenty-five nonwords were chosen for practice trials prior to block 1 and block 9 of the main experiment.

3.4.1.3.1 *Auditory stimuli*

Two native speakers of British English (1 male, 1 female, different from speakers of Experiment 4 and 5) who were unknown to participants recorded all the nonwords in Sets J and K (see Appendix 6). Other settings of the auditory stimuli were identical to Experiment 4.

3.4.1.3.2 *Distractors*

Both the Hear-and-repeat with orthographic distractors and Hear-and-repeat with non-orthographic distractors conditions incorporate the use of distractors. The 4-letter strings distractors in the Hear-and-repeat with orthographic distractors condition were derived from real words. For example, *guab* was derived from ar/*guab/*ly; and *ctua* was derived from intelle/*ctua*/l. All of the stimuli for the orthographic distractors are shown in Appendix 6. The facial picture distractors in the Hear-and-repeat with non-orthographic distractors condition were black- and-white pictures of unknown individuals selected from the Stirling Face Database (http://www.pics.psych.stir.ac.uk). All faces were in a full frontal position.

3.4.1.4 Procedure

Both reaction time (RTs) and accuracy were measured for the experiment. The experiment began with participants signing a consent form. The main experiment consisted of two parts, the learning phase and the final naming block.

3.4.1.4.1 *Main experiment*

Participants were randomly allocated to one of the following conditions with either Set J or K as the training set, 1) hear and repeat, 2) hear and repeat with orthographic distractors, 3) hear and repeat with non-orthographic distractors, and 4) read aloud. The sets were counter-balanced. After 8 blocks of training, participants were asked to read aloud both Set J and K in the final naming block. The untrained set of nonwords was taken as the baseline to assess whether there is any general improvement of novel words reading after the learning phase.

3.4.1.4.2 *Learning phase (Block 1 – 8)*

Participants were randomly allocated in one of the following conditions 1) hear-and-repeat, 2) hear-and-repeat with orthographic distractors, 3) hear-and-repeat with non-orthographic distractors, and 4) read aloud. Each condition will be explained thoroughly in the following section. All stimuli in the training phase were performed using EPrime2 software (PST Inc). No RTs were recorded. The experimenter sat next to the participant to mark any mispronunciation and provided feedback to participants after each block.

Table 3.8. ***Mean and standard deviation of bigram frequency, neighbourhood size and reading speed of the seven-letter nonwords. Range of the 7-letter nonwords reading speed is also included.***

	Nonwords	
	Set J 7-letter	Set K 7-letter
Log bigram frequency		
Mean	3.31	3.32
S.D.	0.09	0.16
Neighbourhood size		
Mean	0.13	0.13
S.D.	0.52	0.35
Phonemes		
Mean	5.93	6.13
S.D.	0.70	0.83
Naming RT (in ms) from Experiment 2 in Chapter 2		
Mean	626	626
S.D.	37.41	37.68
Range	542 - 676	548 - 669
Naming accuracy from Experiment 2 in Chapter 2		
Mean (max = 24)	23.60	23.87
S.D.	0.63	0.35
Range (max = 24)	22 - 24	23-24

Note. RT = Reaction time (naming latency) in ms; S.D. = standard deviation

3.4.1.4.3 *Hear-and-repeat training*

Participants were trained on the phonology of either Set J or K. A typical trial proceeded as follows: A fixation cross appeared in the centre of the computer screen for 1000 ms, then one of the spoken stimulus item was presented at a comfortable listening level over a professional speaker for 3000 ms along with the presentation of a blank black screen. Participants then repeated the item they just heard. 6% of the randomized trials, a red dot would appear in the centre of the screen for 500 ms. Participants had to press the 1st button of the response box as quickly as possible when they saw one while repeating the invented words. Participants were told that the main focus should be on learning the nonwords. There were eight blocks in this learning phase in which participants heard the 15 nonwords for eight times. The order of nonwords varied between blocks and the order of blocks was fixed between participants. Eighteen nonwords were distributed into three practice trials prior to the main experiment to ensure that participants understood the tasks.

Table 3.9. ***Illustration of the distribution of training and test sessions.***

Participant Group	Block 1 – 8 (training)	Set	Block 9
1	Hear and repeat	J	read aloud Sets J & K
2		K	
3	Hear and repeat with orthographic distractors	J	
4		K	
5	Hear and repeat with non-orthographic distractors	J	
6		K	
7	Read aloud	J	
8		K	

3.4.1.4.4 *Hear-and-repeat training with orthographic distractors*

All of the procedure of this condition is exactly the same as those in the hear-and-repeat condition, except that the spoken stimuli were presented along with a sequence of letter strings that changed every 500ms. The letter strings were 4-letter fragments that were chosen from real words. Participants then repeated the nonword they just heard. For 6% of the randomized trials, a red dot would appear in the centre of the screen for 500 ms. Participants had to press the 1st button of the response box as quickly as possible when they saw one while repeating the invented words.

3.4.1.4.5 *Hear-and-repeat training with non-orthographic distractors*

All of the procedure of this condition is exactly the same as those in the hear-and-repeat condition, except that the spoken stimuli were presented along with a sequence of faces that changed every 500ms. *The* participants then repeated the nonword they just heard. For 6% of the randomized trials, a red dot would appear in the centre of the screen for 500 ms. Participants had to press the 1st button of the response box as quickly as possible when they saw one while repeating the invented words.

3.4.1.4.6 *Reading aloud training*

Participants were trained on the orthography of either Set J or K by reading the nonword aloud as quickly and accurately as possible once the nonword appeared on screen. The procedure of read aloud training is exactly the same as the hear-and-repeat training except the spelling of nonwords were shown on screen in lower case Times New Roman font point 18 instead of the blank black screen. Participants would not hear the pronunciation of the nonwords. For 6% of the randomized trials, a red dot would appear in the centre of the screen for 500 ms. Participants had to press the 1st button of the response box as quickly as possible when they saw one while reading the invented words.

3.4.1.4.7 *Testing phase (Block 9)*

After being trained for 8 blocks, either visually or verbally of Set J and K, participants were asked to name both Set J and K in block nine. The participants were tested individually. They were seated approximately 60 cm from the monitor. They wore a set of headphones with a high sensitivity microphone attached. The microphone was linked to a voice key that detected vocal input. Participants were instructed to pronounce the nonwords clearly, and as quickly and accurately as possible, without coughs or hesitations. Before the experimental session, there was a practice session in which 7 items were shown for participants to become familiar with the experimental procedure. It also gave the experimenter the opportunity to adjust the microphone, if necessary.

Reaction times were recorded, using the E-prime software system version 2 (Psychology Software Tools Inc.; www.pst-net.com/eprime). On each trial, a black fixation cross appeared in the centre of the screen for 1000 ms. The fixation cross was followed by a nonword displaced in lower case Times New Roman font point 18, which was presented for 2000 ms. Participants were asked to pronounce the nonword as quickly and accurately as possible once they saw the nonword appeared on screen. After the presentation of the nonword, a blank screen was shown for 1000 ms. Subsequently the next trial started. No feedback was given. The experimenter sat next to the participant to mark any mispronunciation and provided feedback to participants after each block.

3.4.2 Results

3.4.2.1 Red dots catch trials

Hesitations and failures to activate the response box were removed from the analysis along with RTs less than 100 ms or longer than 2.5 SDs above the mean (defined separately for each participant across the eight blocks after removal of the very short RTs). Hesitations and failures to activate the response box occurred on 113 trials (0.2% of the total). An additional 78 RTs were removed at the stage of RT trimming (1.6%), leaving 4801 RTs for analysis (96.2% of the total). Table 3.10 shows the percent of accuracy and RTs for the correct, trimmed red dot catch trials.

3.4.2.1.1 *Catch trials reaction accuracy and RTs*

Non-parametric Kruskal-Wallis test was adapted as accuracy was found to violate the assumption of normality (Kolmogorov-Smirmov test of normality, $p < .05$). Accuracy levels were high (average 95.2% correct across the eight blocks of the four conditions). Ceiling effects meant that there was no significant difference between the reaction accuracy across the four conditions, $H(104) = 5.38$, $p = .146$. There is also no significant difference between the reaction time RTs across all four conditions, $F_1(3, 100) = 1.73$, $MSE = 2363$, $p = .165$, $\eta_p^2 = .049$. This implies that all four conditions require a similar amount of attention.

3.4.2.2 Data trimming

Only RTs for correct responses were analysed. Naming errors, hesitations and failures to activate the voice key were removed from the analysis along with RTs less than 100 ms or longer than 2.5 SDs above the mean (defined separately for each participant for trained and untrained items after removal of the very short RTs). Naming errors, hesitations and failures to activate the voice key occurred on 113 trials (0.9% of the total). An additional 30 RTs were removed at the stage of RT trimming (0.9%), leaving 3060 RTs for analysis (98.1% of the total). Table 3.10 shows the percent of accuracy and RT results for correct, trimmed responses. Accuracy never fell below 97% correct for any condition of trials. The results of the statistical tests are presented in Appendix 2.

3.4.2.3 Naming accuracy

The non-parametric Kruskal-Wallis test was adopted as accuracy was found to violate the assumption of normality. Accuracy levels were high (average 98.8% correct across the trained and untrained items of the four conditions). Ceiling effects meant that there was no significant difference between the trained items across the four conditions, *H*(104) = 2.55, *p* = .466. There was a significant effect in the untrained items, *H*(104) = 11.39, *p* < .01. This was because the accuracy in the hear-and-repeat with non-orthographic distractors condition was slightly lower than the other 3 conditions. The accuracy never fell below 97% correct in the untrained items in any conditions. Post-hoc comparison of Mann-Whitney tests were used to follow up this finding. A Bonferroni correction was applied and so all effects are reported at a .017 level of significance. None of the pairwise comparison between read-aloud to any of the other conditions reached significance, including the comparison of accuracy between the read aloud and the hear-and-repeat conditions, *U*(50) = 259.50, *Z* = 2.30, *p* = .022; the read aloud condition and hear-and-repeat condition with orthographic distractors, *U*(50) = 404, *Z* = 1.84, *p* = .066; the read aloud condition and hear-and-repeat condition with non-orthographic distractors, *U*(50) = 316.50, *Z* = 0.49, *p* = .624. Figure 3.9 shows the mean accuracy of the trained and untrained items in the four conditions.

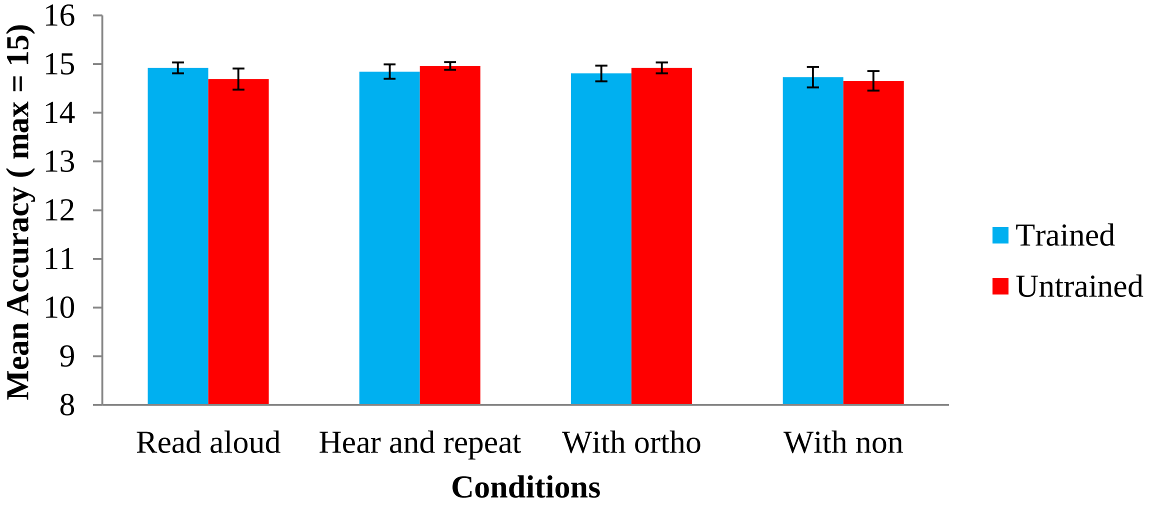

***Figure 3.9.* The mean accuracy of the trained and untrained items in the read aloud, hear-and-repeat, hear-and-repeat with orthographic distractors and hear-and-repeat with non-orthographic distractors conditions. Error bars show 95% confidence intervals.**

*with ortho = hear-and-repeat with orthographic distractors condition, with non = hear-and-repeat with non-orthographic distractors condition.

3.4.2.4 Naming latencies (RTs)

Figure 3.10 shows the pattern of RTs for correct, trimmed responses to repeated (trained) and nonrepeated (untrained) items in the read aloud, hear-and-repeat, hear-and-repeat with orthographic distractors and hear-and-repeat with non-orthographic distractors conditions. Inspection of Figure 3.10 suggests the pattern of the naming RTs was very similar among each condition. The naming RTs of the trained items were always faster compared to the untrained items.

The RT data were first analysed across the trained and untrained items with factors of Training (trained vs untrained), Conditions (all four conditions including read aloud, hear-and-repeat, hear-and-repeat with orthographic distractors, and hear-and-repeat with non-orthographic distractors). There was a significant effect in training (faster overall RTs on trained than untrained items). The two-way interaction between Training and Conditions was also significant. Bonferroni-corrected t-test ($\propto$ = .01) comparing RTs to trained and untrained items in each condition found significant difference in all four conditions. The data were analysed further with the naming RTs difference (untrained – trained item RTs) of each condition using non-parametric tests.

3.4.2.4.1 *RTs difference between conditions*

Figure 3.11 focuses on the naming RTs difference in each condition across trained and untrained items. The read aloud condition had the greatest difference in the naming RTs between trained and untrained items, followed by hear-and-repeat, hear-and-repeat with non-orthographic distractors and hear-and-repeat with orthographic distractors conditions. Table 3.10 shows the details of the RT difference for correct, trimmed responses between trained and untrained items.

Table 3.10. ***Mean latencies of correct, trimmed responses, standard errors (SE), per cent correct responses for 7-letter nonwords and red dot fillers in each condition of Experiment 6.***

	Read aloud		**Hear-and-repeat**		**Hear-and-repeat with orthographic distractors**		**Hear-and-repeat with non-orthographic distractors**	
Red dot fillers								
Mean RT	376		386		390		369	
SD	15.08		9.77		13.57		9.38	
% correct	95.8		97.3		91.2		96.5	
7-letter nonwords	trained	untrained	trained	untrained	trained	untrained	trained	untrained
Mean RT	512	573	531	571	534	563	506	535
SD	68.0	96.4	67.6	74.2	78.5	82.6	87.4	112.5
% correct	99.0	99.7	99.5	97.9	98.7	99.5	98.2	97.7
RTs difference between trained and untrained items								
Mean RT	61		40		29		30	
SD	54.6		29.7		30.4		43.4	

Note. RT = naming reaction times. SD = standard deviation.

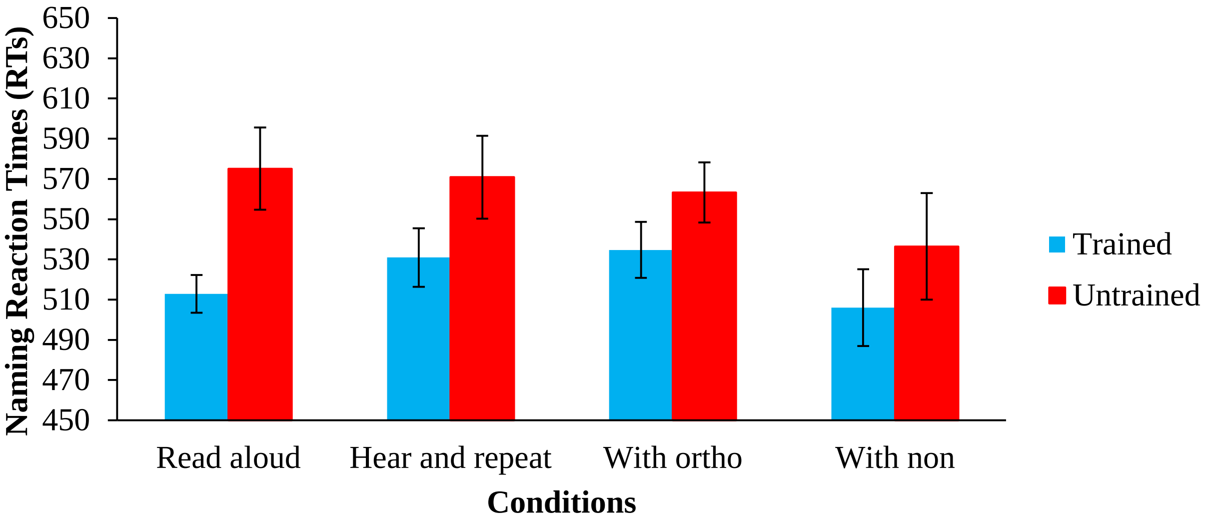

Figure 3.10. **The naming reaction time (RTs) for trained and untrained items in the read aloud, hear-and-repeat, hear-and-repeat with orthographic distractors and hear-and-repeat with non-orthographic distractors conditions. Error bars show 95% confidence intervals. *with ortho = hear-and-repeat with orthographic distractors condition, with non = hear-and-repeat with non-orthographic distractors condition.**

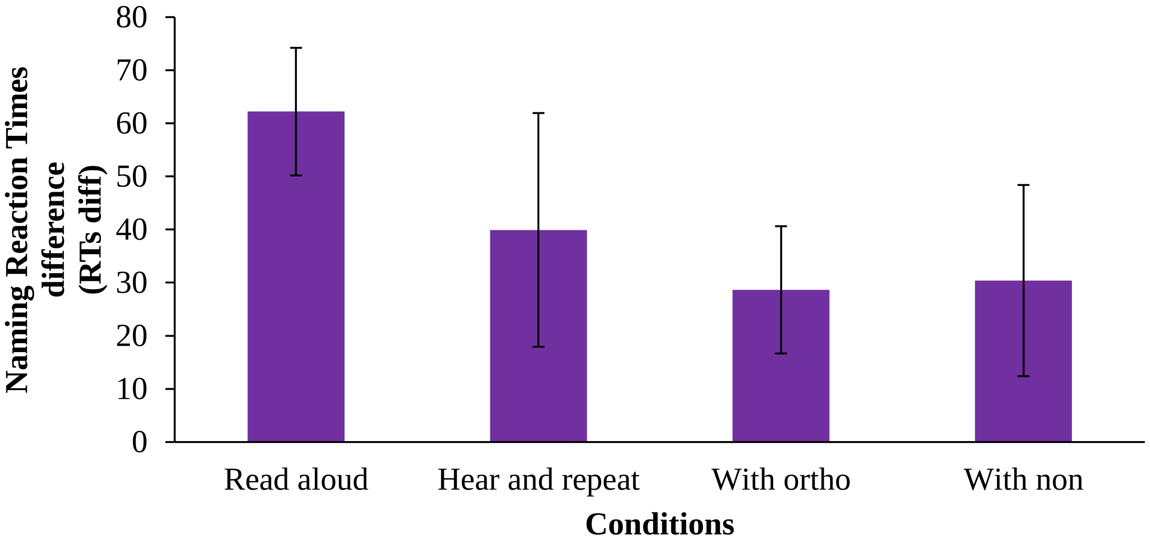

Figure 3.11. **The naming reaction time difference (RTs diff) between trained and untrained items in the read aloud, hear-and-repeat, hear-and-repeat with orthographic distractors and hear-and-repeat with non-orthographic distractors conditions. Error bars show 95% confidence intervals. *with ortho = hear-and-repeat with orthographic distractors condition, with non = hear-and-repeat with non-orthographic distractors condition.**

The non-parametric Kruskal-Wallis test was adopted as naming RTs difference was found to violate the assumption of normality (Kolmogorov-Smirmov test of normality, $p < .05$). The naming RTs difference was analysed with factors of Conditions. There was a significant main effect of Conditions. Post-hoc comparison of Mann-Whitney tests with Bonferroni correction ($\propto = .017$) was used to compare the naming RTs difference of the read aloud and each of the other conditions. There was a significant difference between the read aloud and hear-and-repeat with orthographic distractors condition. The naming RTs difference between the read aloud and hear-and-repeat with non-orthographic distractors condition was also significant.

3.4.3 Discussion

The results for naming latencies in the read aloud and hear-and-repeat conditions were much the same as for the orthographic and phonological training conditions in Experiment 4 and 5. In the context of very high levels of accuracy, all four conditions showed significant faster naming latency to trained compared to untrained items. Focusing on the difference of the naming latency of the trained and untrained items, there was a significant difference between the read aloud and hear-and-repeat with orthographic distractors conditions and the read aloud and hear-and-repeat with non-orthographic distractors conditions.

Putting the result in simple terms, the naming latency differences between training and untrained items was 61 ms for the read aloud condition and 30 ms for the average naming RTs difference for the two hear-and-repeat with distractors conditions. This can be explained that the read aloud condition has provided an extra 30 ms benefit by training participants with the orthography as well as the phonology of the novel words. The high accuracy rate of the red dot demonstrated that participants were indeed paying attention to the distractors on the screen. The result further confirmed that learning is most effective and efficient when both orthography and phonology of the novel words were learned in the read aloud condition. Yet, training the phonology of the novel word in adults is sufficient to build representations in the mental lexicon.

3.5 General Discussion

This chapter aims to explore the role of phonology in visual word learning. The three experiments in the present study yielded much the same pattern of results that nonwords trained for both orthographic and phonology (read aloud condition) was found to be the more efficient and effective compared to training the phonology of the novel words (hear-and-repeat and hear-and-repeat with distractors conditions). Yet, all three experiments also showed that the establishment of a phonological representation of a novel word can be sufficient to result in representations in the mental lexicon even without any encounter with the orthographic form of the novel word. In both Experiment 4 and 5, in the first block of trials, when the nonwords were read for the first time, naming RTs were slow and the difference in RTs between 4- and 7-letter nonwords was obvious. This replicated the result in Chapter 2 and previous studies (e.g. Juphard et al., 2004; Weekes 1997). Averaging over the present three experiments, British native skilled adult readers read the 4- and 7-letter nonwords aloud with mean RTs of 537 ms and 621 ms respectively. An average of 28 ms per additional letter was therefore required in order to read the 7-letter nonwords compared with the 4-letter nonwords in block 1. This result is relatively similar to that observed in Chapter 2.

3.5.1 Is training the phonology of the word sufficient to build representation in the mental lexicon?

The result that learning is most efficient and effective when both orthography and phonology are provided in training matched with the results that were observed from word learning in the developmental and adult literature. As mentioned in section 1.5 in Chapter 1, the self-learning device theory of Share (1995, 1999, 2004) emphasises two mechanisms. Firstly, the strong decoding skills provide young children with a strategy to map a printed word into its spoken form. Secondly, this decoding process then provides a chance for readers to acquire word-specific orthographic information that forms the foundation of efficient word recognition. This theory is in line with developmental studies that suggest training the phonology of novel words poses difficulties for children to acquire lexical entries in the mental lexicon (Ricketts et al, 2009; Rosenthal and Ehri, 2008). Of note is that in both Ricketts et al. (2009) and Rosenthal and Ehri (2008) studies, children were still able to learn the novel words to a certain degree even when they did not encounter the spelling of the novel words. Yet, the learning was 30 – 50% weaker compared to the condition when the orthography of the novel words were shown.

This learning pattern was remarkably similar in adults. As mentioned in section 3.1.4, Maloney et al. (2009) observed the effect of length converged more in the read aloud compared to the case decision condition. Maloney et al. (2009) explained this as that the reading aloud condition required the explicit generation of the phonology of novel words whereas the case decision task only required the implicit generation of the phonology of the novel words. As mentioned in section 1.6 in Chapter 1, Chalmers and Burt (2008) investigated the role of phonological encoding skills in orthographic learning. In the training phase of the study, the orthography of each nonword was presented in the centre of the screen, with (P+) or without its pronunciation (P-). If present, pronunciation began at display onset. Participants were instructed to count the number of consonant clusters in the nonword (to encourage the processing of orthography) and to record their response by key press (m for more than 1, n for not more than 1). They also found that the combined (phonological learning + orthographic learning) condition is significantly better than the condition that was trained on the orthography of the novel words (without the presence of the novel word’s phonology). Yet, simply asking participants to count the number of consonant clusters in the nonword is significant enough to build representation in the mental lexicon as the spelling recognition accuracy in the orthography condition was 73% versus 81% in comparison to the combined learning condition.

These results can be explained by the lexical equality hypothesis. As mentioned in section 3.1.1, Perfetti (1992) suggested that a high-quality lexical representation is complete and accurate in all three components with efficient links between the components. As the read aloud condition has the support of both orthography and phonology whereas hear-and-repeat training mainly has the support of the phonological representations, it is understandable that there will be extra benefit to be trained in the read aloud condition. This leads to the topic of whether there is automatic and mandatory activation of orthographic codes when participants received the hear-and-repeat training which is the topic that I will now turn.

3.5.2 Is there automatic and mandatory activation of orthographic codes of English novel words when participants hear them verbally?

Experiment 4 illustrated that orthographic training is better than phonological training and this effect is more apparent in long items. Experiment 5 demonstrated that orthographic training was truly better than phonological training, not only numerically but also statistically. Yet, there was a possibility that a brief orthographic code was activated when participants listened to the novel words in the phonological training which means that hear-and-repeat training may be an equivalent training of read-aloud condition. Thus, Experiment 6 incorporated distractors in the hear-and-repeat condition with both orthographic and non-orthographic distractors. No significant difference was observed between naming RTs difference between the read aloud and hear-and-repeat condition. Yet, a significant effect was observed when the naming RTs difference was compared to the read aloud and the two hear-and-repeat conditions when orthographic and non-orthographic distractors were incorporated.

As mentioned in section 1.2 in Chapter 1, the orthographic consistency effect has been observed in English and German in a wide range of tasks including naming and lexical decision task (e.g. Ziegler et al., 2001). Yet, whether the same effect can be observed in pseudoword processing has been controversial, with Bürki et al. (2012) and McKague et al. (2008) showing a clear consistency effect for pseudowords, but not Ziegler and Ferrand (1998) or Ventura et al. (2004).

According to the lexical equality hypothesis, it is natural that the feedback consistency effect is more evident from the orthography-to-phonology rather than the phonology-to-orthography pathway. It has been established that phonological information is activated routinely during the visual presentation of words and pseudowords (e.g. Perfetti & Bell, 1991; Unsworth & Pexman, 2003). Yet, the generation of phonology-to-orthography coding would be more difficult as spelling is a more difficult task that required retrieval of a completely specified and accurate orthographic form (Perfetti, 1997).

The result of the current study is in line with Johnston et al. (2004) and McKague et al. (2008) that at least a briefly formed orthographic representation of the novel words were encoded in the hear-and-repeat training prior to the first visual encounter of the novel words. Based on McKague et al. (2008), the orthographic representation that was generated by the phonological training would mainly be framed by the consonants of the novel words. Based on McKague et al.’s (2008) finding, future studies can explore whether there is a benefit from phonological training to reading nonwords that share the same consonants but have different vowels. For example, if McKague et al.’s (2008) suggestion is right, then one would expect that by have phonological training on the nonword *blispod* would reduce the naming RTs of *blespud* as the two nonwords share the same consonants but not the vowels.

Experiment 6 of the current study had attempted to attenuate the automatic and mandatory activation of the orthographic representations in the phonological training by incorporating distractors in the hear-and-repeat condition. However, the equivalent distractor that aims to discourage the automatic activation of phonological representation in orthographic training has not been implemented in the current experiment. Future studies can address this issue by adopting an artificial orthography paradigm that utilized novel characters (Taylor et al., 2011).

3.5.3 Predictors of orthographic learning

3.5.3.1 Predicting initial nonword reading speed

Linear mixed effect modelling found TOWRE PDE and phonological awareness made a significant contribution to predicting 7-letter nonwords naming speed in Block 1 when vocabulary and rapid digit naming were taken into account. This result replicated and extended the literature that found phonological awareness as a crucial predictor of word learning ability (e.g. Muter et al., 2004; Hatcher et al., 1994). Young et al. (2002) extended the result by showing phonological awareness was found to be a significant unique contributor of spelling achievement in adults, and the effect was over and above non-verbal IQ and rapid digit naming skills.

How, theoretically, can the predictive effect of phoneme awareness on nonword reading skills be explained? Perhaps the first interpretation would be in terms of a causal theory that sees learning to read an alphabetic script as critically dependent on a child's possessing adequate phoneme sensitivity. This viewpoint was advocated on the basis of philosophical and clinical observation (e.g., Savin, 1972) long before the vast majority of studies in this area were conducted (e.g., Snowling, 2001). This theory can be explained in two ways, including both proximal and distal factors. Firstly, as a proximal factor, it means that learning to read an alphabetic script requires an explicit awareness of phonemes in speech and the knowledge in which those phonemes are linked to letters (the alphabetic principle; Byrne & Fielding-Barnsley, 1989). Alternatively, as a distal factor, it can mean the phonological awareness task is one way of assessing the quality or integrity of the child's phonological representations that underlie the ability to learn to decode (Snowling & Hulme, 1994). This means that the development of phonological skills in turn has important consequences for other aspects of development, including reading. In the case of reading skills, much of the research has focused on the development of phonological awareness and its role in facilitating the development of reading. This explanation poses phonological awareness task tap a metaphonological level of representation that is itself partly a product of literacy skills (e.g. Muter et al., 2004).

3.5.3.2 Predicting the improvement in orthographic learning

3.5.3.2.1 *Expressive vocabulary*

The ability to learn novel words in both orthographic and phonological training conditions (measured here as the change in RTs to longer nonwords between blocks 1 and 10 of day 1) was predicted by expressive vocabulary. Nation and Snowling (2004) found that vocabulary knowledge accounted for unique variance in children's word reading measured concurrently at 8 years of age and longitudinally when the children's reading was retested at 13 years of age, even when nonword reading and phonological skills were included in the analysis.

Scarborough's (1998) study also found that kindergarten vocabulary skill is associated with later reading performance. Out of the 19 predictors, expressive vocabulary was the significant predictor of later reading skills even when alphabet knowledge, print exposure and story recall skills were controlled. Similarly, Ricketts et al. (2009) found that vocabulary predicted the ability of normal 8 to 10-year-olds to read words with irregular or exceptional spellings but did not predict their ability to read nonwords. Irregular words are words that violate the grapheme-phoneme correspondences of English, e.g. *think* and *know*. This means that children cannot rely on the nonlexical procedures to read these words correctly: they must focus on word-specific learning and the creation of lexical entries. Finally, Braze et al. (2007) also found that vocabulary skills captured a unique variance in adult's reading comprehension ability. These studies are therefore in line with the present findings, albeit the majority of it targeted to a younger group of readers.

The association between vocabulary and word learning can be explained in three ways. Firstly, if one has a larger vocabulary size, novel words they encounter in reading are likely to have more orthographic and phonological neighbors; that is, familiar words that look and sound like the novel words, differing from them by only a few letters or phonemes. Storkel, Armbruster, and Hogan (2006) taught adults novel spoken words paired with novel objects through stories and pictures. Learning was better for nonwords with many neighbors than for nonwords with few neighbors. In the DRC model, words that are already established in the orthographic and phonological lexicons support the processing of new words or nonwords which resemble them. This is conducted through the excitatory and inhibitory interactions between the two lexicons and the systems that encode and represent letter and phoneme sequences. Those interactions allow the model to process nonwords with many neighbors more efficiently than nonwords with fewer neighbors. Lexical support for novel words during learning could explain the advantage for nonwords with many neighbors reported by Storkel et al. (2006) and the benefit of a larger vocabulary found by the aforementioned study and in the present study.

Secondly, as mentioned in section 1.4.3 in Chapter 1, Wise et al. (2007) suggested vocabulary knowledge can help word identification in another two routes. The first route is that vocabulary skills reflect whether there is an effective and efficient connection between the stored phonological representations and the correspondent orthographic patterns. This means that people with good vocabulary skills have a solid phonological representation of words that they can map it to written words. Thirdly, good depth of vocabulary knowledge (expressive vocabulary) may imply that one is faster in encoding, organizing and retrieving of the representations of words. The latter explanation is indirectly associated to Storkel et al. (2006) theory as if one has a larger vocabulary size, there will be more words in the lexicon for one to relate the new words to, which eventually will speed up the encoding process.

3.5.3.2.2 *Rapid digit naming*

The ability to learn novel words in orthographic, but *not* the phonological training condition was also predicted by rapid digit naming skills. Research into the predictive association between rapid naming and reading has found mixed results. When prior reading skill was controlled for, Wagner et al. (1997) did not find a unique variance between rapid digit naming and reading in children. Yet, Wagner et al. (1993) found rapid naming was a separate factor to phonological awareness and memory in a confirmatory factor analysis. Young et al. (2002) also demonstrated that this link between RDN and word identification was also observed in adults that RDN was still a significant predictor to word identification even when phonological awareness and non-verbal IQ was taken into account.

The reason as to why RDN is a significant predictor to orthographic word learning has to be explained by the double deficit hypothesis (Wolf & Bowers, 1999). This theory suggested that phonological deficits (i.e. deficit in decoding) and the processes that underpin reading are two separable sources. As stated by Wolf et al. (2000), orthographic word learning requires a combination of skills including lower perceptual, attentional, articulatory, and lexical retrieval process and higher level cognitive and linguistic processes, and each of these stages demand swift and efficient rates of processing. The result of the current study is in line with the RDN literature that it suggests the link between RDN and word learning may be due to the broad demands of rapid execution in the higher level processes during reading. There is a possibility that adults with higher RDN have a more efficient lexical access between the orthographic input lexicon and phonological output lexicon. This explanation is particularly reasonable in the current study as there is only a significant predictor link between RDN and orthographic learning but not in the phonological learning condition.

3.5.4 Conclusion

The three experiments reported in this chapter found a similar result of those in Chapter 2 that repeatedly presenting novel words (nonwords) to be read aloud as quickly as possible results in 1) a reduction of naming latencies and 2) a reduction in the effect of length. This chapter further demonstrates that by asking participants to hear and repeat novel words can also produce these results to a good extent. This means that while reading aloud is an effective and efficient training to learn new novel words, training the phonology of the novel word is sufficient to build lexical entries in the mental lexicon. While there is a possibility that participants activate the orthographic codes during the hear-and-repeat training, Experiment 6 shows that trained items were still learned better compared to untrained items even when distractors were included in the hear-and-repeat training in order to attenuate the activation of orthographic codes.

In accordance to the ‘lexical quality hypothesis’, learning was better in the read aloud training conditions which participants had the benefit to receive training in both the orthography and the phonology of the stimuli. The fact that the convergence of RTs to shorter and longer nonwords was also shown in the phonological training condition in which participants were only trained on the phonology of the stimuli is in line with previous developmental (Reitsma, 1983) and adult literature (Chalmers & Burt, 2008; Sandak et al., 2004) that phonology is an essential part of learning new words.

Experiment 4 also found that phonological awareness and TOWRE PDE were the significant predictors of the naming latency of the 7-letter nonwords that were read for the first time even when vocabulary and RDN were taken into account. This result replicated and extended the literature that found phonological awareness as a crucial predictor of word learning ability (e.g. Muter et al., 2004; Hatcher et al., 1994). It was also found that expressive vocabulary accounted for the improvement in orthographic and phonological word learning when phonological awareness and Literacy score were included in the analysis. Finally, RDN also accounted for a significant variance in the improvement in orthographic but *not* phonological word learning. This is in line with Wolf's (1997) double deficit hypothesis that phonological awareness and the processes underlying naming speed (RDN) taped into two distinct sources of reading dysfunction.

There are, issues remaining to be resolved, one of which is the equivalent condition that involves orthographic learning that does not involve the activation by the phonological codes. Furthermore, only one task was used in each element that taps into the cognitive skills of the participants. Future studies should incorporate a few more tests in each predictor. Nonetheless, the paradigm developed here can be considered as a tool for understanding the role of phonology in visual word learning. One application would be to study how the effect of length differs in various language groups; for example, Spanish (Ferrand, 2000) and Chinese (Ho & Bryant, 1997). This can lead to further understanding in how the effect of length operates in alphabetic and logographic languages.

4 Reading and lexicalisation in English

4.1 Introduction

When native adult speakers of English read aloud invented, word-like nonwords, the speed with which they can convert those letter sequences into spoken output increases with number of letters in the nonwords. In contrast, when the same native speakers are asked to read aloud familiar words, reaction times (RTs) are faster and the effect of letter length is greatly reduced (Ellis et al., 2009; Mason, 1978; Weekes, 1997; Ziegler, Perry, Jacobs, & Braun, 2001; see also Hogaboam & Perfetti, 1978, for a similar result in children). This chapter utilizes the same learning paradigm that was developed in Chapter 2 to understand how the newly learned items integrate with existing knowledge in the mental lexicon with high- and low-frequency words. Unlike Experiment 3 in Chapter 2, high-frequency and low-frequency words are integrated with nonwords in this experiment and the two sessions are 28 days apart instead in order to understand whether the newly learned words showed good retention over an interval of 4 weeks.

As mentioned in section 1.7 in Chapter 1, a good amount of work has shown that the naming RTs for words is often faster than nonwords (the lexicality effect) and while the effect of length was apparent and highly significant for nonwords, it was declined for low-frequency words and nonsignificant for high-frequency words. Smaller effects of length for high- than low-frequency words have also been reported by Balota et al. (2004), Cosky (1976) and Yap and Balota (2009).The interaction between lexicality and length in French was also observed by Juphard et al. (2004).

4.1.1 How can the DRC model explain the interaction between lexicality and length

The findings of the interaction between lexicality and length in previous literature can be explained by the DRC model (Coltheart et al., 2001). As mentioned in section 1.8 of Chapter 1, when one reads a nonwords (or an unlearned word), there is no other way to read it besides from processing the nonword serially, using a left-to-right manner by operating the grapheme-phoneme correspondence(Coltheart & Rastle, 1994; Coltheart et al., 2001; Rastle et al., 2003; Weekes, 1997). Thus, the length effect of reading aloud is just an inevitable consequence of processing from the non-lexical route. Words, on the contrary, have visual representations in the orthographic input lexicon and phoneme-based representations in the phonological output lexicon which allow the reader to convert the words from print to sound rapidly and lexically. Though the nonlexical route cannot be switched off, but the lexical conversion of high frequency words operates so quickly that a response will be made before the nonlexical route delivers a rule-based pronunciation. This explains why there is little or no interaction between lexical and nonlexical processes in reading aloud high-frequency words.

Lexical conversion of low-frequency words from print to sound is slower and may overlap in time with the delivery of the rule-based pronunciation by the nonlexical route. This provides an opportunity for the interaction between lexical and nonlexical routes in the reading of low-frequency words. An example is words with irregular or exceptional spelling-sound correspondences (e.g. have, said) are read slower compared to words with regular or consistent spelling-sound correspondence (e.g. cat, farm), this regularity effect is also more evident in reading low- than high-frequency words. (e.g., Andrews, 1992; Hino & Lupker, 2000; Jared, 1997; Monaghan & Ellis, 2002; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; see Coltheart et al., 2001, pp. 221-222 & 231-233). The conjunction of the lexical and nonlexical processes within the same time frame for low- but not high-frequency words could explain why the effect of length (commonly thought to be an indicator of nonlexical processing) is larger for low- than high-frequency words (Balota et al., 2004; Cosky, 1976; Weekes, 1997; Yap & Balota, 2009).

Experiment 3 in Chapter 2 and previous literature (e.g. Maloney et al., 2009) already showed that it is possible to use the combination of naming speed and impact of length on naming as the indicators of word learning. By the time RTs in Experiment 3 of Chapter 2 had reached asymptote, naming latencies were similar to those reported by Weekes (1997) for high-frequency words (around 500 ms). It can be appealing to conclude that it only takes 5 or 6 repetitions to an unfamiliar nonword before representations have been created that allow the nonword to be processed with the same level as a high-frequency word. Such a claim could be problematic on its grounds as it compares RTs for nonwords being read for the 6th time within the same session of an experiment to RTs for high frequency words being read for the very first time in an experiment.

4.1.2 Priming effect for real words

There have been many reports suggesting that recognition of familiar English words benefit by repeated presentation (i.e., repetition priming under conditions where each presentation of a word is obviously visible rather than masked). Humphreys, Besner, and Quinlan (1998) showed that only when the stimuli of each presentation was clearly visible produced long-lasting repetition effects). There are inconsistent findings as to whether the effects of repetition are greater for low- than high-frequency words. Low-frequency words have been observed to facilitate more from repetition than high-frequency words in a lexical decision task, no matter whether the prime is clearly visible (Coane & Balota, 2010; Duchek & Neely, 1989; Schilling, Rayner, & Chumbley, 1998) or very brief (Forster & Davis, 1984).

While the frequency effects has been widely reported to be a significant predictor of English word naming (Balota et al., 2004; Brysbaert et al., 2011; Cortese & Schock, 2013; Yap & Balota, 2009), evidence for an interaction between frequency and repetition in word naming is limited. Experiment 3 of Scarborough et al. (1977) showed that while the naming RTs of high-frequency, low-frequency words and nonwords all reduced from the 1st to 2nd presentations, the low-frequency words did not benefit more from the repetition compared to high-frequency words (i.e. the interaction between frequency and repetition was not significant). Balota and Spieler (1999) obtained a significant interaction between frequency and repetition in a lexical decision task, but they only found a significant main effect of frequency (naming RTs of high-frequency were 12 ms faster than low-frequency words overall) for the naming task, the interaction between frequency and repetition was not significant. The significant interaction between frequency and repetitions seems to only be observed in Colombo, Pasini, and

Balota (2006) in an Italian word naming task. Italian has more transparent orthography than English with reliable mappings between spellings and sounds, which means that participants could rely more on the nonlexical route while reading Italian words compared to English. This may contribute to why there is a significant interaction of frequency and repetition in Italian but not in English in previous studies.

4.1.3 The present experiment

The present experiment investigates the process of word learning and the effects of frequency, length and repetition in English. This experiment involved the repeated presentation of interleaved high-frequency words, low-frequency words and nonwords to native speakers of English in two testing sessions 28 days apart. Theoretical interest lies in the relative effects of length on naming latencies for high-frequency words, low-frequency words and nonwords, the extent to which those latencies (RTs) converge for shorter and longer words and nonwords, and the persistence of training/repetition effects over a 28-day retention interval.

4.2 Experiment 7: The process of reading and lexicalisation in English

4.3 Methods

4.3.1 Participants

Participants were 25 undergraduate students of the University of York, UK (13 female, 12 male) with a mean age of 20.08 years (S.D. = 2.68; range 18 - 31). All were native speakers of English with normal or corrected-to-normal vision and no history of reading or language problems. Participants received either course credit or a small payment in return for their participation. The experiment was approved by the Ethics Committee of the Department of Psychology, University of York.

4.3.2 Materials

The experimental stimuli were 24 high-frequency words, 24 low-frequency words and 24 nonwords. Within each set, 12 items contained 4 letters and one syllable while 12 contained 7 letters and two syllables. The short and long high-frequency words, low-frequency words and nonwords were matched on initial letters and phonemes. Twelve different onsets were used to make the items as distinct as possible. To optimize voice key activation, none of the stimuli began with a voiceless fricative (‘f’, ‘s’, ‘sh’ or ‘th’).

Two frequency measures were used in creating the sets of high and low frequency real words – the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993, 1995) which is based on samples of written and spoken English, and SUBTLEX frequencies (Brysbaert & New, 2009) which are based on the subtitles of English films and television programmes. High-frequency words had frequencies of at least 50 occurrences per million words of English on both measures while low frequency words had frequencies below 24 on both measures.

The nonwords were pronounceable letter strings generated by the WordGen program (Duyck et al., 2004) and based on the CELEX and Lexique databases (Baayen et al., 1993, 1995; New et al., 2004). The age of acquisition value (AoA) of each item was gathered from Kuperman, Stadthagen-Gonzalez, and Brysbaert (2012) in which they asked participants to enter the age (in years) at which they thought they had learned the word. Imageability ratings were collected from Bird, Franklin, and Howard (2001). The imageability ratings were made on a 7-point scale (with 1 being the least imageable and 7 begin the most imageable). The mean rating for each item was multiplied by 100 by the authors to give ratings on a scale 100 to 700. The nonwords in the current experiment used a different set from those used by Experiments 1 - 6. The nonwords were matched to the real words on letter length, syllable length, initial letters and phonemes, and mean log bigram frequency from WordGen. None of the words or nonwords was the orthographic or phonological neighbour of any of the other words and nonwords. The details of the experimental stimuli on the matching variables are shown in Table 4.1. All of the experimental stimuli are shown in Appendix 6. Eighteen additional high-frequency words, low-frequency words and nonwords (6 of each) were selected for use in practice trials.

The AoA values of the 4- and 7-letter items in each frequency group were included in an ANOVA analysis with Frequency and length as factors. There was a significant main effect of Frequency, $F_1(1, 11) = 12.02$, $MSE = 31.38$, $p < .005$, $\eta_p^2 = .522$, with the high-frequency words having a lower AoA value (mean = 5.92, S.D. = 0.84) than the low-frequency words (mean = 7.54, S.D. = 1.77).There was no significant effect of Length, $F_1(1, 11) = 2.27$, $MSE = 9.71$, $p = .160$, $\eta_p^2 = .171$, and the interaction between Frequency and Length, $F_1(1, 11) = 1.42$, $MSE = 2.35$, $p = .258$, $\eta_p^2 = .114$. The same analysis was conducted for the imageability values. A significant effect was not found for the main effect of Frequency, $F_1(1, 11) = 2.51$, $MSE = 13838$, $p = .142$, $\eta_p^2 = .186$; Length, $F_1(1, 11) = 0.01$, $MSE = 17.52$, $p = .968$, $\eta_p^2 = .000$, and the interaction between Frequency and Length, $F_1(1, 11) = 0.40$, $MSE = 8829$, $p = .538$, $\eta_p^2 = .035$.

4.3.3 Procedure

Participants were tested individually. After completing a consent form, participants were given practice on the task. This involved reading 18 items (6 high-frequency words, 6 low-frequency words and 6 nonwords, with half the items of each type containing 4 letters and half 7 letters). The experiment then began with the 72 stimuli being presented in a random order in block 1 (12 short high frequency words, 12 long high frequency words, 12 short low frequency words, 12 long low frequency words, 12 short nonwords and 12 long nonwords).

Table 4.1. ***Details of the stimuli used in Experiment 8.***

	4 letters			7 letters		
	High frequency	Low frequency	Nonwords	High frequency	Low frequency	Nonwords
CELEX word frequency						
Mean	169	11.8	–	179	9.8	–
S.D.	76	5.9	–	182	7.3	–
SUBTLEX word frequency						
Mean	142	7.6	–	157	6.4	–
S.D.	59	4.9	–	152	5.2	–
Age of acquisition						
Mean	5.69	6.87	–	6.15	8.21	–
S.D.	0.69	2.0	–	1.51	2.62	–
Imageability						
Mean	428	489	–	454	461	–
S.D.	98	126	–	119	113	–
Mean log bigram frequency						
Mean	3.38	3.32	3.38	3.38	3.30	3.39
S.D.	0.15	0.22	0.12	0.10	0.19	0.09

**Note:* S.D. = standard deviation

The procedure was the same as those used in Experiment 1 of Chapter 2, except the 72 stimuli were presented once in a random order. Participants were informed when the block was complete and pressed the space bar on a computer keyboard to initiate the next block when they were ready to continue. This process was repeated across 10 blocks with the stimuli being presented in a different random order in each block. The experimenter noted any trials in which the participant misread a nonword, hesitated or made a false start or other form of error. No feedback was given at any point. Participants returned 28 days later for a second session which repeated the practice items and the 10 blocks of experimental stimuli.

4.4 Result

Only RTs for correct responses were analysed. Naming errors, hesitations and failures to activate the voice key were removed from the analysis along with RTs less than 100 ms or longer than 2.5 SDs above the mean (defined separately for each participant in each block and for each length after removal of the very short RTs). Naming errors, hesitations and failures to activate the voice key occurred on 80 trials (0.2% of the total). An additional 40 RTs were removed at the stage of RT trimming (0.1%), leaving 35,880 RTs for analysis (99.6% of the total). Table 4.2 shows the accuracy and RT results for correct, trimmed responses.

4.4.1 Accuracy

The non-parametric Friedman test was adopted as accuracy was found to violate the assumption of normality (Kolmogorov-Smirmov test of normality, $p < .05$). Accuracy was very high (99.8% correct overall and never less than 98% correct in any condition or block of trials). Given the high levels of accuracy in all the conditions, the nonparametric Friedman test found no significant difference in the accuracy to high-frequency words, low-frequency words and nonwords across days 1 and 28 together, $\chi^2(2) = 0.67$, $p = .717$. Wilcoxon matched pairs, signed ranks tests found no significant difference between accuracy for 4- vs 7-letter nonwords across the two sessions for high-frequency words, $W(25) = 15.0$, $Z = 4.50$, $p = .317$; and nonwords, $W(25) = 2.50$, $Z = 1.73$, $p = .0.84$ (4-letter mean = 11.98, S.D. = 0.05; 7-letter mean = 11.96, S.D. = 0.07). There was a marginal significant effect of length for 4- vs 7-letter nonwords across the two sessions for low-frequency words $W(25) = 10.0$, $Z = 1.89$, $p = .059$ (4-letter mean = 11.96, S.D. = 0.07; 7-letter mean = 11.98, S.D. = 0.05). Wilcoxon matched pairs, signed ranks tests also found a significant difference for the overall accuracy on day 1 and day 28, $W(25) = 49.5$, $Z = 2.26$, $p < .05$, with Day 28 having a slightly higher accuracy rate (mean = 12.00, S.D. = 0.01) compared to Day 1 (mean = 11.65, S.D. = 0.11). Figure 4.1 shows the mean accuracy of the 4- and 7-letter items in each condition.

4.4.2 Naming latency

Figure 4.2 shows the pattern of RTs for correct, trimmed responses to high-frequency words, low-frequency words and nonwords across days and blocks. RTs were analysed using ANOVA and when Mauchly's test of sphericity was significant, the Greenhouse-Geiger correction was applied. A series of analyses were performed to address different questions. Full details of the statistical analyses are shown in Appendix 3 where effect sizes are reported in terms of the partial eta squared statistic (η_p^2). The main findings will be summarized here.

4.4.2.1 Day 1 Block 1

Analysis of the RTs to high-frequency words, low-frequency words and nonwords in block 1 of day 1 allows a direct comparison with the results obtained by Weekes (1997) and the many other studies that have compared naming RTs to high- and low-frequency words, or to words and nonwords, presented just once. Inspection of Figure 4.2 and Table 4.2 suggests that RTs were fastest to high-frequency words in block 1 of day 1, slower for low-frequency words and slowest for nonwords, and that the effect of length on RTs was largest for nonwords, smaller for low- frequency words and smallest for high-frequency words. Day 1, block 1 RTs were analysed with Stimulus type and Length as factors. There was a highly significant main effect of Stimulus type, with fastest overall RTs to high frequency words (mean = 520 ms) followed by low-frequency words (mean = 535 ms) then nonwords (mean = 589 ms). The main effect of Length was highly significant, with faster RTs overall to 4- than 7-letter stimuli. The Stimulus type x Length interaction was also highly significant, reflecting the fact that effects of length on naming latencies in block 1 of day 1 were largest for nonwords (mean length effect = 84 ms), smaller for low-frequency words (mean length effect = 17 ms) and smallest for high-frequency words (mean length effect = 8 ms).

Separate analyses compared high- with low-frequency words, high-frequency words with nonwords and low-frequency words with nonwords on block 1 of day 1. The comparison of high- with low-frequency words produced a significant effect of Stimulus type, with RTs being faster to high- than low-frequency words. The main effect of Length was also significant, but the interaction between Stimulus type and Length was not significant.

The comparisons of high-frequency words with nonwords, and low-frequency words with nonwords, found significant main effects of Stimulus type (faster RTs to words than nonwords in both analyses) and Length (faster RTs to 4- than 7-letter items). The interaction between Stimulus type and Length was also significant, reflecting the fact that length effects were greater for nonwords than for either high- or low-frequency words.

Bonferroni-corrected *t*-tests were used to compare RTs to 4- and 7-letter stimuli for high-frequency words, low-frequency words and nonwords in block 1 of day 1. The length effect for high-frequency words was not significant, the effect for low-frequency words was marginally significant, while the effect for nonwords was highly significant (see Appendix 3).

In sum, the statistical analysis of RTs in block 1 of day 1 supported the indications in Figure 4.2 and Table 4.2. High-frequency words were read aloud more rapidly than low-frequency words which were, in turn, read more rapidly than nonwords. The effect of length was nonsignificant for high-frequency words, weak (at best) for low-frequency words, and highly significant for nonwords.

4.4.2.2 Day 1 blocks 1-10

The next set of analyses addressed the changes in RTs on day 1 across blocks 1 to 10. Figure 4.2 and Table 4.2 suggest that RTs to all three types of stimulus decreased with repetition across blocks, that any length effects found for words in block 1 rapidly disappeared across subsequent blocks, and that the large length effect for nonwords in block 1 reduced across later blocks but did not disappear completely.

Table 4.2. ***Mean RTs (with S.D.) and accuracy (with S.D. and percent of correct) for four- and seven-letter high-frequency words, low-frequency words and nonwords presented on Day 1 and Day 28.***

	Day 1										Day 28									
Blocks	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
High-frequency words																				
4-letter																				
Mean RTs	516	495	489	492	485	478	480	476	476	470	494	471	475	480	482	478	484	485	483	481
S.D.	62	71	68	70	72	69	64	63	61	67	65	57	66	63	70	72	67	70	67	73
Mean Acc	11.9	12.0	12.0	12.0	11.9	12.0	11.9	12.0	12.0	11.9	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0
S.D.	0.44	0.00	0.20	0.20	0.40	0.20	0.44	0.00	0.00	0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20
% correct	99.0	100	99.7	99.7	99.3	99.7	99.0	100	100	99.3	100	100	100	100	100	100	100	100	100	99.6
7-letter																				
Mean RTs	524	500	492	486	478	475	480	477	471	471	493	478	475	473	477	481	484	481	475	475
S.D.	68	70	73	62	64	60	69	67	60	62	69	67	65	63	65	72	78	68	81	76
Mean Acc	11.9	12.0	12.0	12.0	11.9	12.0	12.0	12.0	11.8	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0
S.D.	0.28	0.20	0.00	0.00	0.44	0.00	0.00	0.20	0.47	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% correct	99.3	99.7	100	100	99.0	100	100	99.7	98.7	100	100	100	100	100	100	100	100	100	100	100

	Day 1										Day 28									
Blocks	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Low-frequency words																				
4-letter																				
Mean RTs	526	503	499	497	485	485	480	484	485	484	494	481	485	488	483	501	489	484	486	485
S.D.	65	76	70	72	66	66	69	67	66	69	57	62	60	71	73	87	75	65	78	68
Mean Acc	11.8	12.0	12.0	11.9	12.0	11.9	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0
S.D.	0.52	0.20	0.20	0.40	0.00	0.28	0.20	0.20	0.20	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% correct	98.0	99.7	99.7	99.3	100	99.3	99.7	99.7	99.7	99.7	100	100	100	100	100	100	100	100	100	100
7-letter																				
Mean RTs	543	503	497	496	485	480	483	485	485	476	493	485	487	494	482	491	481	478	483	472
S.D.	88	81	74	71	66	74	63	67	67	59	66	70	72	78	72	79	65	59	80	77
Mean Acc	11.9	12.0	12.0	12.0	12.0	11.9	12.0	11.9	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0
S.D.	0.33	0.00	0.20	0.00	0.20	0.44	0.00	0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% correct	99.0	100	99.7	100	99.7	99.0	100	99.3	100	100	100	100	100	100	100	100	100	100	100	100

	Day 1										Day 28									
Blocks	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Nonwords																				
4-letter																				
Mean RTs	547	510	498	501	494	492	499	493	488	474	499	489	490	491	482	498	497	492	494	482
S.D.	95	72	76	82	70	71	74	72	67	57	56	68	68	74	59	80	75	69	70	72
Mean Acc	11.8	12.0	12.0	11.9	11.9	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0
S.D.	0.47	0.20	0.00	0.28	0.28	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% correct	98.7	99.7	100	99.3	99.3	100	100	100	100	99.7	100	100	100	100	100	100	100	100	100	100
7-letter																				
Mean RTs	631	549	534	527	526	516	504	515	507	497	545	501	501	498	499	496	497	505	491	487
S.D.	139	106	104	85	103	85	90	80	86	85	96	68	67	63	77	76	73	75	69	78
Mean Acc	11.9	12.0	11.9	12.0	12.0	12.0	12.0	11.8	12.0	12.0	11.9	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0
S.D.	0.33	0.20	0.28	0.20	0.20	0.20	0.20	0.47	0.20	0.00	0.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
% correct	99.0	99.7	99.3	99.7	99.7	99.7	99.7	98.7	99.7	100	99.0	100	100	100	100	100	100	100	100	100

**Note:* S.D. = standard deviation; RTs = naming reaction times; Acc = naming accuracy

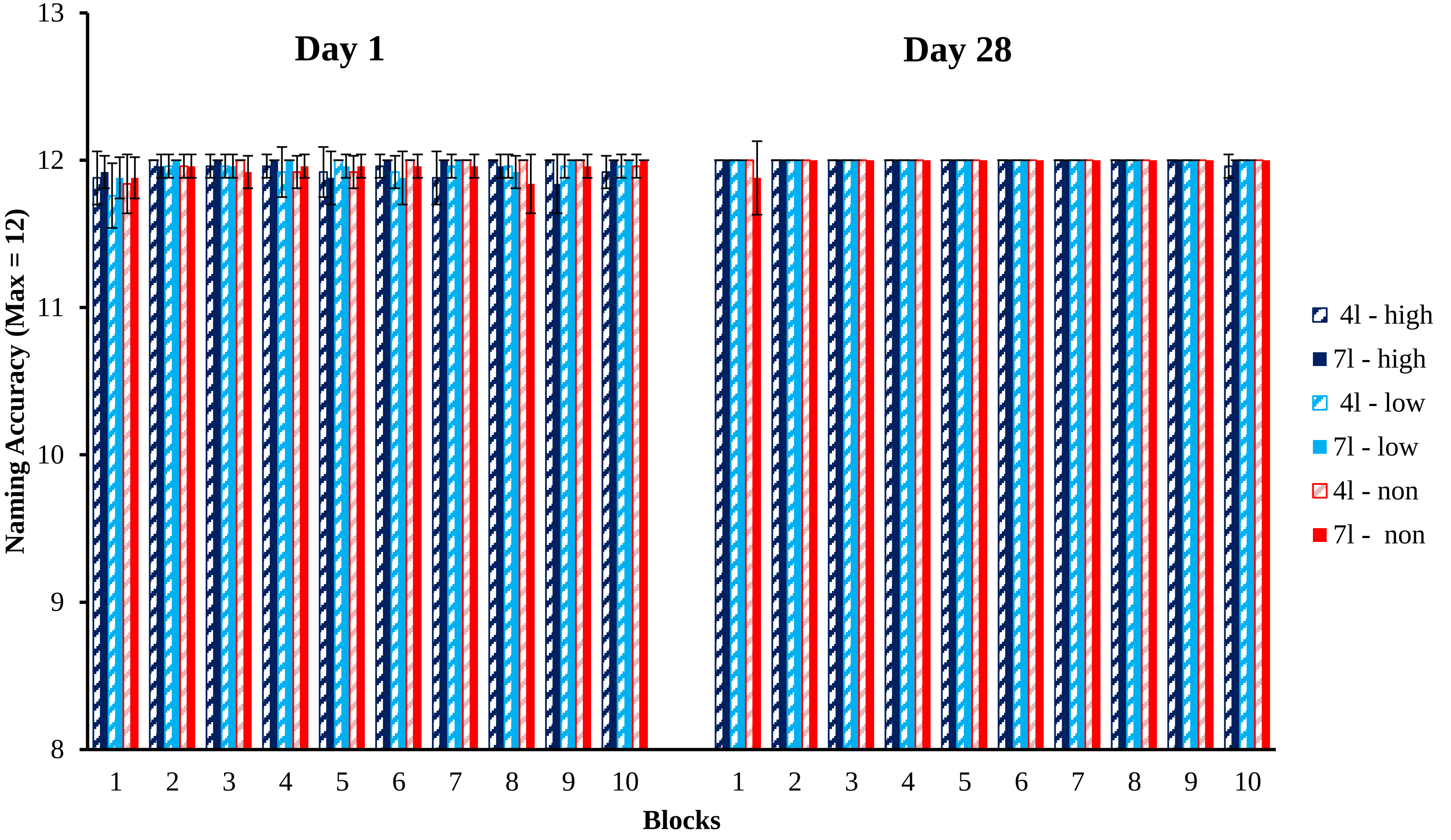

Figure 4.1. **The accuracy of naming 4- and 7-letter high-frequency words, low-frequency words and nonwords in Blocks 1 to 10 in Days 1 and 28. Error bars show 95% confidence intervals and their absence means that no variance was present in the data.**

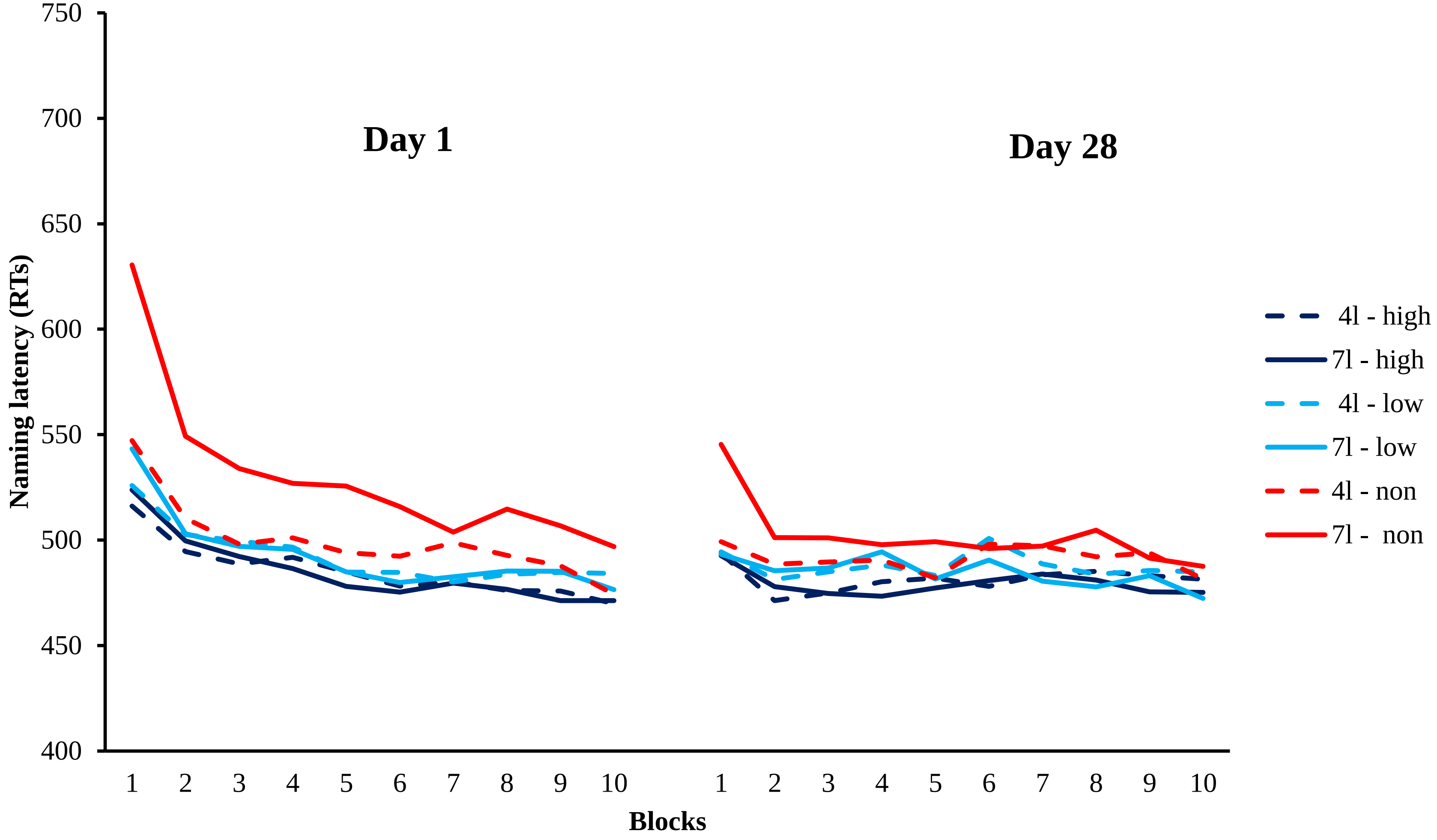

Figure 4.2. **The naming latency (RTs) for 4- and 7-letter high-frequency words, low-frequency words and nonwords in Blocks 1 to 10 in Days 1 and 28.**

Day 1 RTs were analysed with Blocks, Stimulus type and Length as factors. The full results are shown in Appendix 3. All of the main effects and interactions were significant. Follow-on analyses compared RTs across the 10 blocks to high- and low-frequency words, high-frequency words and nonwords, and low-frequency words and nonwords. The analysis comparing high- with low-frequency words found a significant effect of Blocks, with RTs decreasing between blocks 1 and 10, and a significant effect of Stimulus type, with faster overall RTs to high- than low-frequency words. There was no overall effect of Length and none of the interactions was significant.

The analyses comparing high-frequency words with nonwords, and low-frequency words with nonwords produced similar results. The main effects of Blocks, Stimulus type and Length were all significant: overall RTs decreased across blocks, were faster to both high- and low-frequency words than to nonwords, and were faster to shorter than longer items. Stimulus type interacted with Length (greater length effect for nonwords than for either high- or low-frequency words). Blocks interacted significantly with Stimulus type (greater decrease in RTs across blocks for nonwords than words). Blocks also interacted with Length (greater decrease in RTs across blocks for 7- compared with 4-letter stimuli). The three-way interaction between Blocks, Stimulus type and Length was also significant in both of the comparisons between words and nonwords.

In view of the significant three-way interaction between Blocks, Stimulus type and Length, RTs were analysed separately for high-frequency words, low-frequency words and nonwords. There were significant main effects of Blocks for all three types of stimulus (decline in RTs across blocks). There was no overall effect of Length for either high- or low-frequency words and no interaction between Blocks and Length for either set of words. Nonwords, in contrast, showed a significant main effect of Length (faster overall RTs to 4- than 7-letter nonwords) and a significant Blocks x Length interaction (reduction in the length effect for nonwords across blocks). Bonferroni-corrected *t*-tests found differences between RTs to 4- and 7-letter nonwords that were significant in blocks 1 to 6, marginally significant (by the demanding standards of Bonferroni-corrected *t*-tests) in blocks 5, 8 and 10, and nonsignificant in blocks 7 and 9.

In sum, RTs to high-frequency words, low-frequency words and nonwords on day 1 all decreased across blocks as a result of repeated presentations, asymptoting around block 5 or 6. The difference in RTs to high- and low-frequency words that was apparent in the analysis of block 1 remained significant across day 1. Overall RTs to nonwords were slower than to either high- or low-frequency words, particularly for longer stimuli. RTs decreased more across blocks for nonwords than for either high- or low-frequency words. Unlike the real words, RTs to nonwords showed a reduction in the size of the length effect with repetitions though convergence between RTs to shorter and longer nonwords did not occur convincingly on day 1.

4.4.2.3 Retention of information across 28 days: comparison of day 1 block 10 with day 28 block 1.

Figure 4.2 suggests some slowing of RTs between the end of day 1 and the beginning of the second testing session on day 28, most noticeably for the longer nonwords. Changes in RTs over the 28-day retention interval were analysed by comparing RTs in block 10 of day 1 with RTs in block 1 of day 28 with Delay (day 1 block 10 vs. day 28 block 1), Stimulus Type and Length as factors. Of interest are the significant main effect of Delay (slower overall RTs in block 1 of day 28 than block 10 of day 1), the significant Delay x Stimulus type interaction and the marginally significant Delay x Stimulus type x Length interaction, indicating that delay had different effects on shorter and longer words and nonwords. Bonferroni-corrected *t*-tests compared RTs in day 1 block 10 with RTs in day 28 block 1 separately for 4- and 7-letter words and nonwords. While the increase in RTs to both 4- and 7-letter nonwords across the 28-day retention interval was significant, the changes in RTs for high- and low-frequency words were, at most, only marginally significant.

In sum, RTs to nonwords slowed between the last block of day 1 and the first block of day 28 while RTs to both high- and low-frequency words changed little or not at all. Comparison of RTs at the beginning of day 1 and the beginning of day 28 (Figure 4.2 and Table 4.2) indicates that the effects of repetition on naming of all stimulus types were retained to a considerable extent across the 4-week retention period.

4.4.2.4 Day 28 blocks 1-10.

Figure 4.2 and Table 4.2 suggest that RTs on day 28 recovered quickly from any slowing down between the end of day 1 and the start of day 28. There is no clear difference between RTs to high- and low-frequency words in the later blocks of day 28 and no obvious effects of length. RTs to the longer nonwords fell markedly from block 1 to block 2. In the later blocks of day 28, RTs to nonwords approached those for real words and the difference in RTs to longer and shorter nonwords virtually disappeared.

RTs on day 28 were analysed in the same way as for day 1, beginning with an overall analysis with Blocks, Stimulus type and Length as factors. The main effect of Stimulus type was significant, but unlike day 1, the main effects of Blocks and Length were not significant. All the interactions were, however, significant. The Blocks x Stimulus type interaction indicated a greater change across blocks on day 28 for nonwords than words, the Blocks x Length interaction reflected a greater change across blocks for 7- than 4-letter stimuli, and the Stimulus type x Length interaction reflected a greater overall length effect for nonwords than words. The significant three-way Blocks x Stimulus type x Length interaction reflected the fact that the change in RTs across the early blocks on day 28 was greatest for the longer nonwords.

Follow-on analyses compared day 28 RTs to high- and low-frequency words, high-frequency words and nonwords, and low-frequency words and nonwords. In the comparison of high- and low-frequency words, the only significant effect was the three-way interaction between Blocks, Stimulus type and Length, reflecting block-to-block fluctuations in the magnitude of length effects and of the difference between RTs to high- and low-frequency words.

The comparisons of high- and low-frequency words to nonwords produced broadly similar results. In both analyses the main effect of Blocks was significant and interacted with Stimulus type (larger change across blocks for nonwords than words). The effect of Length also interacted with Blocks (greater change in the length effect across blocks for nonwords than words). Significant three-way interactions between Blocks, Stimulus type and Length reflected the fact that the greatest change in RTs was for the longer nonwords between blocks 1 and 2 of day 28.

Separate analyses of day 28 RTs to high-frequency words, low-frequency words and nonwords with Blocks and Length as factors found no significant change in RTs across blocks and no length effect for either high- or low-frequency words. RTs to nonwords showed effects of Blocks and Length combined with a significant Blocks x Length interaction. Bonferroni-corrected *t*-tests comparing RTs to 4- and 7-letter nonwords across blocks 1-10 of day 28 found a significant effect of length in block 1 only.

In sum, there was no overall difference between RTs to high- and low-frequency words on day 28 and no consistent effects of length or change across blocks. RTs to nonwords, which had shown the greatest increase between the end of day 1 and the start of day 28, decreased between blocks 1 and 2 (especially RTs to the longer nonwords). The length effect for nonwords was only significant in block 1 (though numerical convergence between RTs to shorter and longer nonwords only occurred around block 6 of day 28.

4.4.2.5 Overall analysis of day 1 and day 28 RTs

The separate analyses of RTs in days 1 and 28 suggested different patterns of performance on the two days. Those indications were evaluated using a global analysis of RTs across both days with Days, Blocks, Stimulus type and Length as factors. The emphasis was on interactions involving Days as a factor. As with the other analyses, the full results are shown in Appendix 3.

Day interacted with Blocks (larger overall change across blocks on day 1 than day 28), Stimulus type (larger difference between stimulus types on day 1 than day 28) and Length (larger overall effect of length on day 1 than day 28). Those interactions were modified by significant three-way interactions between Day, Blocks and Stimulus type (greater difference in the effect of blocks across stimulus types on day 1 than day 28) and Day, Stimulus type and Length (stronger interaction between stimulus type and length on day 1 than day 28).

In sum, the global analysis supported the indications in the separate analyses of days 1 and 28 that RTs changed more across blocks on day 1 than day 28, that differences in RTs to high-frequency words, low-frequency words and nonwords were greater on day 1 than day 28, and that the effects of length were reduced on day 28 compared with day 1, particularly for nonwords in the later blocks of day 28.

4.5 Discussion

Novel words are essentially new words when participants first see it. This chapter utilizes the same learning paradigm that was developed in Chapter 2 to understand how the newly learned items integrate with existing knowledge in the mental lexicon with high- and low-frequency words. The two testing sessions were 28 days apart. This experiment showed that learning of new novel words was completed in the first learning session. The learned materials were then retained for 28 days with no extra revision. Thus, the learning effect that was observed in the nonwords can be seen as an interaction of AoA (late acquired by adult participants) and frequency (high in recent exposure). Naming RTs were faster to high- than low-frequency words and faster to words than nonwords. Those differences were larger for longer items and diminished across blocks and days as items were repeated. As mentioned in Experiment 2 of Chapter 2, other work has shown that RTs to easier and more difficult items are more similar when the different types of items are presented in separate blocks than when they are interleaved (e.g. Lupker et al., 1997; Rastle et al., 2003). In the current experiment words and nonwords were randomly interleaved: it is possible that the frequency and lexicality effects that were observed here (and the length effects) would have been greater if the different types of stimulus had been blocked rather than interleaved. When comparing the naming RTs of the nonwords in day 1 block 1 to those in Experiment 3 of Chapter 2, the nonwords naming RTs of the current experiment were faster (7-letter nonwords RTs was 613ms in the current experiment and 703ms in Experiment of Chapter 2).

4.5.1 Word learning and the DRC model

In line with the findings of the literature (e.g. Balota et al., 2004; Weekes; 1997; Jupard et al., 2004; Zieger et al., 2001), the result of the current experiment found a significant length effect of nonwords in day 1 block 1, a marginally significant length effect of low-frequency words and a non-significant effect for high-frequency words. The result supports the notion that nonwords require processing on the grapheme-phoneme correspondence route. Though low-frequency words had already built representations in the orthographic input lexicon and the phonological output lexicon, given that lexical conversion of low-frequency words from print to sound is slow and may overlap in time with the delivery of the rule-based pronunciation by the nonlexical route, a marginal effect of length was observed when participants read the low-frequency words for the first time. Finally, though the nonlexical route cannot be switched off, the lexical conversion of high frequency words operates so quickly that a response was made before the nonlexical route delivers a rule-based pronunciation. This explains why a length effect was not found for the high-frequency words for the current experiment.

As mentioned in Chapter 3, reading novel words aloud includes the component of both orthography and phonology. In DRC terms, new representations of the novel words would be created in both the orthographic input lexicon and the phonological output lexicon in the lexical route. Naming RTs of nonwords diminished across the first 6 blocks on day 1 before reaching asymptote. This implies it took around 6 blocks for participants to build representations in the orthographic and phonological lexicons. This result is consistent with the findings of those observed in Experiment 1, 2, and 3 in

Chapter 2. The result of the current experiment is also in line with the literature that children (Ehri & Saltmarsh, 1995; Hogaboam & Perfetti, 1978; Reitsma, 1983; Share, 1999) and adults (Gaskell & Dumay, 2003; Maloney et al., 2009; Salasoo et al., 1985) showed rapid word learning.

4.5.2 Lexicalization of new words and priming of words

RTs of both high- and low-frequency words also improved through blocks in day 1. This can be explained by a repetition priming effect. Stark and McClelland (2000) used an implicit learning task (CID) to understand the repetition priming effect. During the CID task, participants are alternately presented with a letter string and a mask, with the duration of the letter string presentation increasing within each fixed-length cycle. In this way, as the trial progresses, the stimulus appears to become clearer. The participant's task is to identify the letter string by verbally naming the letters as soon as possible while maintaining accurate responses. In Experiment 1 of Stark and McClelland (2000), words, pseudowords (word-like nonwords) and nonwords (letter strings that looked unlike real words) were mixed in the presentation of the CID task. Not only did they find a significant effect of the stimulus type, with words being identified faster than pseudowords and pseudowords were recognized faster than nonwords; they also found a significant main effect of repetitions and an interaction between stimulus type and repetition, with pseudowords having a larger priming effect than words. The significant main effect of stimulus type and repetition and the interaction between the two factors were consistent with the result of the current experiment.

4.5.3 Frequency effect for familiar words are modulated by recent experience

On day 1 of the current experiment, RTs to nonwords were similar in block 2 (7-letter naming RTs = 549ms) to those of low-frequency words in block 1 (7-letter naming RTs = 543ms). Nonwords RTs in block 4 (7-letter naming RTs = 527ms) were similar to those of high-frequency words in block 1 (7-letter naming RTs = 524ms). But RTs to words declined with repetitions too. Even by the end of day 28 there was still a difference between RTs to words and nonwords. In the present experiment, the 20 presentations of the nonwords across two sessions four weeks apart was not sufficient to lexicalise the nonwords to the point where performance on them was indistinguishable from either the high- or the low-frequency words. This leads to the question as to how much of the frequency effect is due to recent experience (repetitions and recency effect) and how much to lifespan frequency?

Scarborough et al. (1977) are the first study to look into how the impact of recent exposure affects the cumulative frequency effect. Cumulative frequency refers to the total number of times that a word is encountered in an individual’s lifetime. In Experiment 2 of Scarborough et al. (1977), participants had to make lexical decision judgement on high-, low-frequency words and nonwords. Words and nonwords were interleaved and were presented up to 3 repetitions. They had also found an interaction of frequency and repetition, with low-frequency words RTs reduced more through repetitions. The authors then concluded this result showed that the recency and cumulative word-frequency effects are closely tied. While these two components are highly linked, the authors suggested that the recency effect is only a part of the word-frequency effect. This claim is in line with the result of the current experiment. In block 1 day 1, the length effect was 8ms for high-frequency words and 17ms for low-frequency words; in block 10 day 28, the length effect was 6ms for high-frequency words and 13 ms for low-frequency words. Thus, the recent exposure of 20 presentations helped to reduce the length effect of the low-frequency words by 1/3. Yet, the cumulative frequency may contribute to the fact that even after 20 presentations, the length effect of high-frequency is still half that of the low-frequency words. Of note is that the high- and low-frequency words in the current experiment also differed on AoA values. Thus, there is a possibility that it is AoA that contributes to the difference of length effect in high- and low-frequency words. Future studies can use the same learning paradigm as this experiment but train participants on different sets of words that vary on frequency, but not AoA or imageability.

Colombo et al. (2006) tried to tease apart the contribution of word cumulative frequency and familiarity in a lexical decision task. The authors suggested that one way to produce changes in familiarity of a word is through repetition of the word within the experimental context. In a regression analysis, both word frequency and familiarity, but not imageability, significantly predicted the RTs of lexical decision response. Two classes of explanations of repetition priming effects are suggested. The first assumes a temporary modification to the process of lexical access. As a result of recent activation, the lexical representation of a word is left in a state of increased accessibility (Forbach, Stanners, & Hochhaus, 1974). The second explanation is based on the idea that the first presentation of a word establishes an episodic memory trace that is contacted when the same word is presented again (Feustel, Shiffrin, & Salasoo, 1983b; Logan, 1990). The result of the current experiment would favour the first explanation more as the word learning in this experiment is lexical learning rather than episodic learning.

4.5.4 Memory retention of learned materials

The current experiment showed some increase in RTs at the start of day 28 for the English nonwords but the benefits of repetition were reinstated in block 2 of day 28. It only took participants one block to read as quickly as in block 10 in day 1. Even with the increase for nonwords across the delay, RTs at the start of day 28 were faster than at the start of day 1. This is in line with the literature that was mentioned in section 1.6 of Chapter 1 that once adults learn the new words, they often retain the information for a long period of time, even up to a year (Breitenstein et al., 2004; Salasoo et al., 1985; Tamminen & Gaskell, 2008). Similar to the result of the current experiment, a warm-up effect was observed in Salasoo et al.'s study, in which participants showed lower accuracy of learned pseudowords before they re-familiarized themselves with the learned letter strings again a year after they learned the new items.

The result of the current experiment is in line with the literature that the effect of repetition priming is quite long-lasting. By using an enumeration task to ask participants to judge whether the number of letters in a given letter string was odd or even, Hauptmann and Karni (2002) showed that the repetition priming effect saturates after the 6th presentation (same as the learning effect of the current experiment), and it lasted until 44 hours after the initial training session. By using an object naming task, Wiggs, Weisberg, and Martin (2006) had also shown that the priming effect in adults is long-lasting, even when a 1-month intervenes between the initial experience and subsequent priming test of an item. Though there was a decrease of priming effect that occurred after a week of the initial training session, the objects that were named in the initial session were still named significantly faster than the filler items in the final naming session that was 1 month apart.

4.5.5 Conclusion and future direction

The current experiment found a similar result of those in Chapter 2 and 3 that repeatedly presenting novel words to be read aloud as quickly as possible results in 1) a reduction of naming RTs and 2) a reduction in the length effect. This chapter further demonstrates that by interleaving high-, low-frequency and nonwords in 20 blocks, both high- and low-frequency words also benefit from repetition priming in that both of their naming RTs reduced. The current experiment also confirms that the frequency effect for familiar words is modulated by recent experience. There is also good retention of the learned novel words even when the two sessions were 28 days apart. Though the naming RTs were a bit slower in day 28 block 1, it only took participants one block before they could read the words and nonwords as quickly as at the end of day 1.

Future research could attempt to understand whether the process of building representations in the orthographic and phonological lexicons will be different for people with poorer phonological skills, e.g. adults with dyslexia and second language speakers. In the current experiment, high- and low-frequency words were not matched on AoA. Future research can try to separate these factors, establishing for example if length effects are greater for late than early acquired words. It will be helpful for future studies to distinguish the effects of letter length from the effects of neighbourhood size as well.

5 Visual word learning in adults with dyslexia

5.1 Introduction

The problems that dyslexic children and adults experience in reading and spelling have been well documented, though there is continuing debate about the underlying causes of those difficulties (Snowling, 2001; Van den Broeck & Geudens, 2012; Vellutino, Fletcher, Snowling, & Scanlon, 2004). One aspect of reading skill that has received less attention than most in the literature, nonetheless, is how adults with dyslexia learn new written words and how their ability to learn new words compares with that of normal readers (de Jong & Messbauer, 2011; Ehri & Saltmarsh, 1995; Mayringer & Wimmer, 2000; Reitsma, 1983; Share & Shalev, 2004; Thomson & Goswami, 2010). The current chapter utilizes the same implicit learning paradigm that was developed in Chapter 2 and applies it to understanding visual word learning in groups of dyslexic adults and normally-reading controls.

As children grow older, reading becomes an important source of new words which they must learn to recognize and understand if they are to function effectively (Cunningham, 2006; Nation, 2008; 2009). Nowhere is this more important than in higher education where, if students are to progress satisfactorily, they must learn new words connected with their academic studies that are often encountered first in written form (Mortimore & Crozier, 2006). Higher Education Statistics Agency (HESA) revealed 23,625 dyslexic students domiciled in the UK during 2005/2006. This represents 2.6% of the total higher education population. Thus, studying how adults with dyslexia learn new words is of practical importance. The present study is not concern with how dyslexics learn to relate new words with meanings, but rather with the process by which initially unfamiliar words become familiar through exposure and repetition, reaching the point where they can be recognized and processed as whole units rather than in incremental fashion.

Based on the result of previous chapters, this chapter compares the performance of university and college students with a diagnosis of dyslexia with typically-reading controls on the same task. As in Experiments 3 in Chapter 1, nonwords composed of either 4 or 7 letters were presented 10 times in a first testing session, then 10 more times in a second testing session 7 days later. Accuracy of reading the nonwords aloud was assessed along with naming latencies. Bruck (1990) and Ben-Dror, Pollatsek, and Scarpati (1991) found slower and less accurate reading of both words in nonwords in American college dyslexics than controls. Similar results have been reported for Polish (Reid, Szczerbinski, Iskierka-Kasperek, & Hansen, 2007) and Swedish (Wolff, 2009) dyslexic university students and controls. Less accurate reading aloud of both words and nonwords by student dyslexics than controls was reported by Snowling, Nation, Moxham, Gallagher, and Frith (1997) and Hatcher et al. (2002) in very similar participant groups to those reported here (see also Callens, Tops, & Brysbaert, 2012; Deacon, Cook, & Parrila, 2012). These observations, combined with reports of less proficient reading of both nonwords and words by dyslexic children (Paizi, De Luca, Zoccolotti, & Burani, 2013; Reid et al., 2007; Wolff, 2009; Zoccolotti et al., 2005), leads to the hypothesis that the dyslexic students in this experiment would be slower and possibly less accurate than controls throughout the experiment, not only when the nonwords were presented for the first time, but also after multiple encounters.

It is also expected that the adult dyslexics would show stronger effects of letter length on reading speed than the controls. There are two reasons why such a difference could arise. First, it has often been proposed that nonword reading presents a prominent problem for dyslexics (Herrmann, Matyas & Pratt, 2006; Rack , Snowling & Olson, 1992; though see Van den Broeck & Geudens, 2012). Wimmer (1996), for example, found that 10-year-old German dyslexic children read nonwords more slowly than younger normal readers who were matched to the dyslexics on the speed of reading familiar, high-frequency words. If nonlexical reading is indeed differentially poor in many dyslexics, length effects should be greater in dyslexics than typical readers because the dyslexics will require more time per additional letter to convert that letter into sound.

Second, if dyslexics are slower than typical readers to create new lexical entries, then in the process of an experiment involving 20 presentations of each nonword across two separate sessions, the dyslexics may be slower than the controls to create orthographic and phonological representations for the novel words. The result would be that they spend more time reading nonlexically (with consequent length effects) and would be slower to process the novel words as a whole unit (with reduced length effects). There does not seem to have been any studies of word learning in dyslexia that have involved adult participants, but research involving dyslexic children suggests problems learning both the spoken and the written forms of new words. Regarding the learning of spoken word-forms, Mayringer and Wimmer (2000) found that German-speaking dyslexic children were impaired at learning novel spoken words that were taught as the names of children shown in pictures. In contrast, the dyslexics were unimpaired at learning to associate familiar German names with pictures of children. The authors concluded from this that the dyslexic children's difficulty lay in learning the new spoken words rather than in associating names with people (see also Elbro & Jensen, 2005; Thomson & Goswami, 2010).

Mayringer and Wimmer (2000) suggested that if dyslexics have problems learning new written words, part of those problems could lie in learning the spoken (phonological) forms rather than their written (orthographic) forms. Visual word learning involves creating phonological as well as orthographic representations: difficulties in learning spoken word-forms would be expected to have a knock-on effect on visual word learning.

The few published studies of visual (rather than spoken) word learning in dyslexia suggest, however, that dyslexics have problems learning new written word-forms over and above any problems they experience in learning spoken words (de Jong & Messbauer, 2011; Ehri & Saltmarsh, 1995; O’Brien, Van Orden, & Pennington, 2013; Reitsma, 1983; Share & Shalev, 2004). Reitsma (1983; Experiment 3) compared visual word learning in Dutch children with reading disabilities with learning in a group of younger normal readers. The children first practiced reading aloud novel words embedded in sentences. Three days later, they were asked to read aloud the novel words as quickly as possible as they were presented individually on a computer screen. Half of the novel words were presented in exactly the same written form as in the training while the other half were presented in a form that had a different spelling but was pronounced the same. (An equivalent English example might be to train children to read *breet* then test them three days later on either *breet* or *breat*.) The normal readers were faster to read aloud the versions of the novel words that they had been trained on three days earlier than the re-spelled version, though they were faster on both than on entirely new and untrained nonwords (so faster on *breet* than *breat* but faster on both of them than on *broat*). In contrast, the children with reading disability read both forms of the trained novel words (*breet* and *breat*) faster than the untrained items (*broat*) but showed no difference between the versions of the trained items that preserved the original spellings (*breet*) and the versions that changed those spellings (*breat*). The implication of these results is that the normal readers learned both the orthographic and phonological forms of the novel words in training and retained that knowledge through to the test three days later. The disabled readers remembered at least partial of the phonological forms of the trained novel items across the retention interval but seemed not to retain any detectable orthographic information.

If dyslexic children combine less efficient nonlexical reading with slower creation of lexical entries, it would be expected that they would show a larger length effect in nonword reading than typically-reading controls. It would also be expected that dyslexics would demonstrate larger effects of letter length in word reading arising from the fact that they are less efficient than controls at moving on from processing a new word serially to reading it as a whole unit (lexically). This prediction is supported by reports of stronger effects of letter length on naming latencies for real words in dyslexic children than controls in English, Dutch, German, Spanish and Italian (e.g. Davies, Rodríguez-Ferreiro, Suárez, & Cuetos, 2013; Marinus & de Jong, 2010; Martelli et al., 2014; Paizi, Zoccolotti, & Burani, 2011; Ziegler, Perry, Ma-Wyatt, Ladner, & Schulte-Korne, 2003).

Dyslexics may have difficulty learning new spoken and written word-forms but dyslexic Italian children have been reported to read words faster than nonwords (Paizi et al., 2013) therefore demonstrating some acquisition of word-specific knowledge. Paizi et al. (2013) also reported faster reading of high- than low-frequency words in dyslexic Italian children, indicating that regular exposure facilitates the creation of effective lexical entries in those readers. If dyslexics are capable of building up a vocabulary of words they can read in a relatively holistic manner, albeit more slowly and effortfully than typical readers, that could explain the reduction in the impact of letter length on word reading with age that Zoccolotti et al. (2005) and De Luca, Barca, Burani, and Zoccolotti (2008) observed in both dyslexic Italian children and controls.

Hence, on the basis of this admittedly incomplete literature, much of which is concerned with children rather than adults, it is expected to see at least some signs of word learning in the dyslexic participants in this experiment (i.e., faster naming latencies across blocks and a reduction in the impact of letter length with repeated exposure). It is expected, however, that word learning would occur more slowly in the dyslexic participants than in controls (typical readers) and that if convergence between reading speeds for shorter and longer items was achieved, it would require more presentations of the nonwords.

Finally, the participants were given short battery of tests to characterize their broader cognitive abilities. The cognitive profiles of dyslexic students at the same institution as many of the participants in the present study (the University of York, UK) were described a decade ago by Hatcher et al. (2002) and more recently by Warmington, Stothard, and Snowling (2013b). Hatcher et al. (2002) found that the student with dyslexia performed at comparable levels to normally-reading controls on nonverbal ability (Raven's Advanced Progressive Matrices) but more poorly on a range of measures including verbal ability (WAIS-R vocabulary), word reading and spelling, forward and backward digit span, phonological tasks (object naming, digit naming and spoonerisms [exchanging sounds between words]) and mental arithmetic. Similar profiles were reported by Snowling et al. (1997) and Warmington et al. (2013b) for UK student dyslexics and Callens et al. (2012) for Belgian dyslexic students. A wider review and meta-analysis of dyslexia in adults is provided by Swanson and Hsieh (2009).

In addition to comparing the dyslexics and controls on the test battery, this study used regression analyses to explore the ability of performance on the different cognitive tests to predict two aspects of performance in the experiment, namely initial reading speeds for the longer (7-letter) nonwords and the change in reading speeds across the 10 presentations in the first testing session. Initial reading speeds assess efficiency of converting unfamiliar letter sequences into sounds (in DRC terms, the efficiency of the nonlexical route), while the reduction in reaction times across repetitions assesses the efficiency of word learning and the change from processing dominantly from nonlexical to lexical reading. Previous research has associated the speed and accuracy of reading nonwords or unfamiliar words with phonological awareness (Durand, Hulme, Larkin, & Snowling, 2005; Melby-Lervåg et al., 2012). For instance, Pennington (1990) documented persisting deficits in phonological awareness in adult dyslexics that were particularly linked to problems with nonword reading. Training studies have suggested, however, that phonological awareness must be linked to an awareness of how letters map onto phonemes if improvements in phonological awareness are to be translated into improvements in reading (Hatcher et al., 1994; Melby-Lervåg et al., 2012).

Word learning has been more strongly associated with working memory than with phonological awareness (Avons, Wragg, Cupples, & Lovegrove, 1998; Gathercole, Hitch, & Martin, 1997; Gathercole et al., 1999). For example, Gathercole et al. (1999) reported an association between phonological working memory and vocabulary size in both 4-year-old and teenage children. Experimental studies by Jarrold, Thorn, and Stephens (2009) and Majerus and Boukebza (2013) reported a relationship between verbal working memory and ability to learn the phonological form (rather than the semantic referent) of new words by children and teenagers while Martin and Ellis (2012) found that word learning in an artificial second language by university students was predicted by performance on phonological short-term / working memory tasks. Short-term and working memory have consistently been found to be impaired in dyslexia which may linked to the problems in word learning mentioned above (Swanson, Zheng, & Jerman, 2009). Based on the findings of these studies, this chapter expanded the predictors of cognitive skills that were mentioned in Chapter 3 and included the measure of working memory, word reading (WRAT), spelling, non-verbal skills and motor speed in the current experiment.

5.2 Experiment 8: Visual word learning in adults with dyslexia

5.2.1 Method

5.2.2 Participants

Participants were 30 students with a diagnosis of dyslexia (20 female, 10 male) and 30 typical readers who served as a control group (12 female, 18 male). The dyslexic students had a mean age of 21.5 years (S.D. = 3.6; range 17 - 36) while the controls had a mean age of 20.7 years (S.D. = 3.2; range 17 - 32). All were native speakers of English with normal or corrected-to-normal vision. The participants were students at the University of York (n = 27 per group), York Saint John University (n = 1 per group) and York College (n = 2 per group). The participants with dyslexia had all been diagnosed by a registered educational psychologist and supplied a copy of their diagnosis documents to the experimenters. Individuals with additional learning disabilities, a history of mental illness, epilepsy or other neurological disorders were excluded. Participants received either course credit or a small payment. The experiment was approved by the Ethics Committee of the Department of Psychology, University of York.

5.2.3 Test Battery

The psychological test battery given to all the participants contained tests assessing vocabulary, reading and spelling, phonological awareness, working memory, nonverbal ability and motor speed. Published tests were scored according to the test manuals and the results are presented as standardized scores. The cognitive test battery of vocabulary, word reading, nonword reading, and phonological awareness was the same as those that was used in Experiment 4 of Chapter 3. The additional predictors of the current experiment included 1) spelling skills, 2) working memory, 3) non-verbal and 4) motor speed.

5.2.3.1 Word spelling

This was assessed using the Spelling Subtest of the WRAT 4 (Wilkinson & Robertson, 2006), which requires participants to write single words to dictation.

5.2.3.2 Word reading

In addition to the word reading test of TOWRE (Torgesen et al., 1999), word reading was also assessed using the reading subtest of the Wide Range Achievement Test which involves reading aloud single words of increasing length and difficulty (from see to synecdoche).

5.2.3.3 Working memory

This was assessed using four tests from the Automated Working Memory Assessment (AWMA; Alloway, 2007). All the tests used span procedures in which sequence lengths were increased to the point where three or more errors were made within a block of trials. Standardized scores were calculated for each test.

5.2.3.3.1 *Verbal short-term memory*

This was measured using immediate serial recall of lists of digits presented auditorily at a rate of 1 / sec.

5.2.3.3.2 *Verbal working memory*

This was assessed using a test in which participants were presented with a sequence of spoken sentences. They were required to decide whether each sentence was true or false then recall the final words of each of the sentences at the end of the sequence.

5.2.3.3.3 *Visuospatial short-term memory*

This was assessed using a dot matrix task in which a sequence of red dots appeared in squares of a 4x4 grid at a rate of one per 2 sec. At the end of the sequence, the participant was required to touch the squares of the grid in the same order.

5.2.3.3.4 *Visuospatial working memory*

This was measured using a spatial recall task. Participants were presented with pairs of shapes. The shape on the right always had a red dot in it. The shape on the left was either the same as the one on the right or different. The shape on the left could also be rotated with respect to the one on the right. The participant's task was first to say whether the two shapes were the same or different. After making those judgments to a sequence of pairs of shapes, the participant then had to indicate in the correct order where the red dot was positioned in each of the shapes on the right using a compass display with three points.

5.2.3.4 Nonverbal ability

This was assessed using the matrix reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler , 1999).

5.2.3.5 Motor speed

This was assessed using a set of tapping tasks (Warmington, Hitch, & Gathercole, 2013a). Participants were asked to tap keys on a computer keyboard as many times as possible within 5 seconds. The start and end of each time interval was signaled both visually and auditory. The task consisted of three conditions with 6 trials in each condition. In Condition 1, the participants tapped one key using the index finger of their preferred hand as many times as possible. In Condition 2, the participants alternately tapped two keys using the index finger of their preferred hand as many times as possible. In Condition 3, the participants alternately tapped two keys using the first two fingers of their preferred hand as many times as possible. The score is the average time between taps across the three conditions.

5.2.4 Stimuli and procedure

Set E (see Appendix 6) was used in this experiment. The stimuli and procedure were identical to those mentioned in Experiment 3 of Chapter 2.

5.3 Result

5.3.1 Performance on the test battery

Table 5.1 shows the results for the dyslexics and controls on the battery of tests together with the results of t-tests comparing the two groups along with the effect sizes (*r*; Field, 2009). Dyslexics performed significantly less well than the controls on every test except nonverbal reasoning. The effect sizes for the differences between the groups were largest for nonword reading, followed by spelling and word reading. The effect sizes for the differences between groups on verbal and visuospatial working memory tasks were similar.

5.3.2 Performance of the experimental task

Only RTs for correct responses were analyzed. Naming errors, hesitations and failures to activate the voice key accounted for 200 trials (0.7% of the total). RTs shorter than 200 ms or longer than mean plus 2.5 SDs in each block for each length group were regarded as outliers and removed from the analyses of accuracy and RTs. This led to the loss of a further 445 RTs (1.5% of the total), leaving 28155 RTs (97.8% of the total) for analysis. Table 5.2 shows the mean accuracy and RTs for correct, trimmed responses for both the dyslexic and control groups. The results of the statistical analyses are presented in Appendix 4.

5.3.2.1 Accuracy

Accuracy was very high (97.3% correct overall and never less than 95.5% correct for either group in any condition or block of trials). Given the high levels of accuracy in both groups, nonparametric Mann-Whitney U tests found no significant difference between dyslexics and typical readers on overall accuracy across the two days for either 4-letter nonwords, $U(60) = 464$, $Z = .208$, $p = .835$, or 7-letter nonwords, $U(60) = 346$, $Z = -1.548$, $p = .122$. Wilcoxon matched pairs, signed ranks tests found no difference between accuracy for 4- vs 7-letter nonwords across the two sessions for both groups of participants combined, $W(12) = 23.0$, $Z = 1.26$, $p = .209$. Figure 5.1 shows the mean accuracy of the typical adults and Figure 5.2 shows the mean accuracy of the dyslexic group.

5.3.2.2 Naming latency

The main analyses focused on the RT data from the experimental task. Figure 5.3. Naming RTs to 4- and 7-letter nonwords in dyslexics and controls across two sessions (10 blocks per session). Error bars show 95% CIs.shows the pattern of RTs for correct, trimmed responses across blocks for the dyslexics (in red) and the controls (in blue). Inspection of Figure 5.3 indicates that naming latencies were slower for the dyslexics than the controls throughout the experiment. At the start of the experiment, both groups were slower to read aloud 7- than 4-letter nonwords. The difference in naming RTs for shorter and longer nonwords reduced with repetitions, but the dyslexic participants appear to have required more exposures to the nonwords before the RTs for shorter and longer items converged. These indications were explored in a series of ANOVAs. When Mauchly's test of sphericity was significant, the Greenhouse-Geiger correction was applied. Full details of the statistical analyses are presented in Appendix 4 where effect sizes are reported in terms of the partial eta squared statistic (η_p^2). The important outcomes will be summarized here.

5.3.2.2.1 *Global analysis*

The first ANOVA was a global analysis conducted on the RT data for both testing sessions with Group, Day, Blocks and Length as factors. There were significant main effects of Group (faster overall RTs for the controls than the dyslexics), Day (faster RTs on day 7 than day 1), Blocks (RTs becoming faster across blocks) and Length (faster overall RTs to 4- than 7-letter nonwords). All of the interactions were significant, including the interaction between Group and Length (larger length effects in the dyslexics than the controls) and Groups x Blocks x Length (the reduction in the length effect across blocks occurring more quickly in the controls than in the dyslexics). These results were explored further by means of separate analyses of RTs in day 1 and day 7, including separate analyses of the performance of the dyslexic and control groups on each day.

Table 5.1. ***Results of the dyslexic and typical readers on the psychological test battery.***

	Dyslexics		Typical readers		
	Mean	**S.D.**	**Mean**	**S.D.**	***t tests and effect sizes (r)***
Vocabulary					
WASI Vocabulary	56.5	7.7	63.7	6.8	$t_{(58)} = 3.87$, $p < .001$; $r = .45$
Word reading					
WRAT 4 Reading	99.0	7.4	117.3	12.8	$t_{(58)} = 6.77$, $p < .001$; $r = .66$
TOWRE-SWE	82.0	11.0	97.4	10.7	$t_{(58)} = 7.21$, $p < .001$; $r = .69$
Nonword reading					
TOWRE-PDE	86.6	10.2	108.1	7.7	$t_{(58)} = 12.01$, $p < .001$; $r = .84$
Spelling					
WRAT 4 Spelling	96.5	12.4	121.3	11.9	$t_{(58)} = 7.95$, $p < .001$; $r = .72$
Phonological awareness					
CTOPP Elision	7.3	1.8	9.0	1.7	$t_{(58)} = 3.40$, $p < .001$; $r = .41$
Working memory					
AWMA verbal STM	87.7	12.8	101.5	14.5	$t_{(58)} = 3.92$, $p < .001$; $r = .46$
AWMA verbal WM	93.0	13.9	106.0	14.6	$t_{(58)} = 3.53$, $p = .001$; $r = .42$
AWMA visuospatial STM	90.3	11.6	108.8	13.1	$t_{(58)} = 5.81$, $p < .001$; $r = .61$
AWMA visuospatial WM	95.9	16.1	106.9	11.9	$t_{(58)} = 3.02$, $p < .01$; $r = .37$
Nonverbal ability					
WASI Matrix reasoning	54.6	7.8	55.8	5.7	$t_{(58)} = 0.66$, $p = .510$; $r = .09$
Motor speed	267.7	55.5	224.3	35.1	$t_{(58)} = -3.54$, $p = .001$; $r = .42$

**Note:* S.D. = standard deviation

Table 5.2. ***Mean latencies of correct, trimmed responses, standard deviations (S.D.), and per cent correct responses for 4- and 7-letter nonwords in blocks 1 to 10 of day 1 and day 7 in dyslexics and typical readers.***

	Day 1									
Blocks	1	2	3	4	5	6	7	8	9	10
Dyslexic readers										
4-letter nonwords										
Mean RT	803	729	701	650	646	636	639	629	606	613
S.D.	180	172	151	125	137	115	106	111	117	111
Mean Acc.	12.00	11.87	11.83	11.80	11.67	11.77	11.87	11.67	11.77	11.70
S.D.	0.00	0.35	0.38	0.41	0.61	0.63	0.35	0.48	0.43	0.65
% correct	100.0	98.9	98.6	98.3	97.2	98.1	98.9	97.2	98.1	97.5
7-letter nonwords										
Mean RT	975	838	765	719	702	670	689	660	654	649
S.D.	226	193	161	154	151	132	134	127	117	118
Mean Acc.	11.80	11.87	11.47	11.63	11.63	11.57	11.53	11.73	11.70	11.77
S.D.	0.41	0.35	0.63	0.56	0.56	0.57	0.57	0.45	0.47	0.43
% correct	98.3	98.9	95.6	96.9	96.9	96.4	96.1	97.8	97.5	98.1
Typical readers										
4-letter nonwords										
Mean RT	597	551	529	537	530	530	525	520	512	511
S.D.	91	85	71	89	92	111	93	84	83	109
Mean Acc.	11.97	11.63	11.77	11.80	11.73	11.77	11.57	11.63	11.63	11.77
S.D.	0.18	0.61	0.50	0.48	0.52	0.43	0.63	0.61	0.61	0.50
% correct	99.7	96.9	98.1	98.3	97.8	98.1	96.4	96.9	96.9	98.1
7-letter nonwords										
Mean RT	666	585	568	548	552	539	541	530	528	526
S.D.	126	102	101	99	104	113	93	77	93	113
Mean Acc.	11.47	11.63	11.83	11.60	11.70	11.67	11.77	11.73	11.83	11.77
S.D.	0.51	0.49	0.38	0.50	0.47	0.48	0.43	0.52	0.38	0.50
% correct	95.6	96.9	98.6	96.7	97.5	97.2	98.1	97.8	98.6	98.1
	Day 7									
Blocks	1	2	3	4	5	6	7	8	9	10
Dyslexic readers										
4-letter nonwords										
Mean RT	657	589	586	594	575	587	565	573	572	561
S.D.	166	130	144	151	101	119	101	99	93	106
Mean Acc.	11.80	11.77	11.80	11.73	11.83	11.80	11.63	11.70	11.87	11.77
S.D.	0.41	0.50	0.41	0.45	0.38	0.41	0.56	0.47	0.35	0.43
% correct	98.3	98.1	98.3	97.8	98.6	98.3	96.9	97.5	98.9	98.1
7-letter nonwords										
Mean RT	721	635	618	599	593	582	585	589	593	574
S.D.	167	142	133	144	108	96	108	113	100	104
Mean Acc.	11.73	11.77	11.73	11.70	11.63	11.60	11.63	11.63	11.50	11.63
S.D.	0.45	0.43	0.45	0.53	0.49	0.56	0.49	0.56	0.57	0.49
% correct	97.8	98.1	97.8	97.5	96.9	96.7	96.9	96.9	95.8	96.9

Typical readers										
4-letter nonwords										
Mean RT	525	508	505	491	496	498	492	481	486	484
S.D.	87	85	90	73	79	83	87	85	93	82
Mean Acc.	11.80	11.90	11.80	11.63	11.83	11.80	11.77	11.73	11.77	11.93
S.D.	0.41	0.31	0.41	0.56	0.38	0.41	0.43	0.45	0.50	0.25
% correct	98.3	99.2	98.3	96.9	98.6	98.3	98.1	97.8	98.1	99.4
7-letter nonwords										
Mean RT	562	515	509	503	497	498	492	494	498	488
S.D.	104	82	100	70	66	76	80	77	81	89
Mean Acc.	11.80	11.70	11.73	11.83	11.77	11.57	11.70	11.77	11.73	11.77
S.D.	0.48	0.53	0.45	0.38	0.50	0.50	0.53	0.50	0.45	0.43
% correct	98.3	97.5	97.8	98.6	98.1	96.4	97.5	98.1	97.8	98.1

Note. RT = reaction time (naming latency); Acc. = naming accuracy; S.D. = standard deviation

5.3.2.2.2 *Day 1.*

Day 1 RTs were analyzed with Group, Blocks and Length as factors. There were significant main effects of Group (faster RTs in the controls than the dyslexics), Blocks (RTs becoming faster across blocks) and Length (faster RTs to 4- than 7-letter nonwords). All of the interactions were significant. Day 1 RTs were then analyzed separately for controls and dyslexics. The controls showed significant main effects of Blocks and Length with a Blocks x Length interaction. Bonferroni-corrected t-tests were used to compare RTs to 4- and 7-letter nonwords in blocks 1 to 10. The effect of length was significant for the controls in blocks 1, 2 and 3 but was no longer significant from block 4 onwards. The dyslexics also showed effects of Blocks and Length combined with a Blocks x Length interaction. In their case, Bonferroni-corrected t-tests found effects of length in blocks 1-5, 7, 9 and 10 with marginally significant effects in blocks 6 and 8 (see Appendix 4).

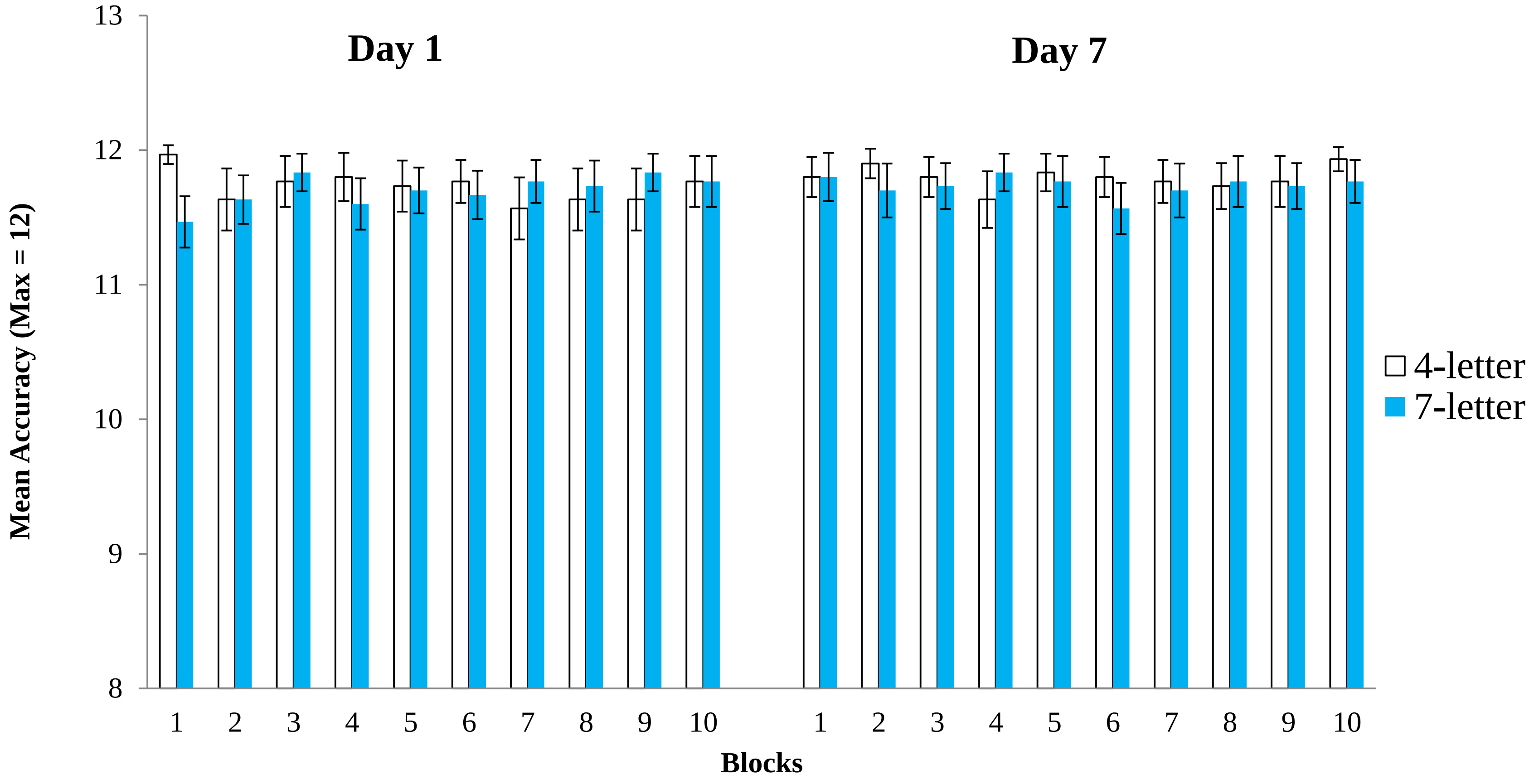

***Figure 5.1*. Naming accuracy to 4- and 7-letter nonwords in controls across two sessions (10 blocks per session). Error bars show 95% CIs.**

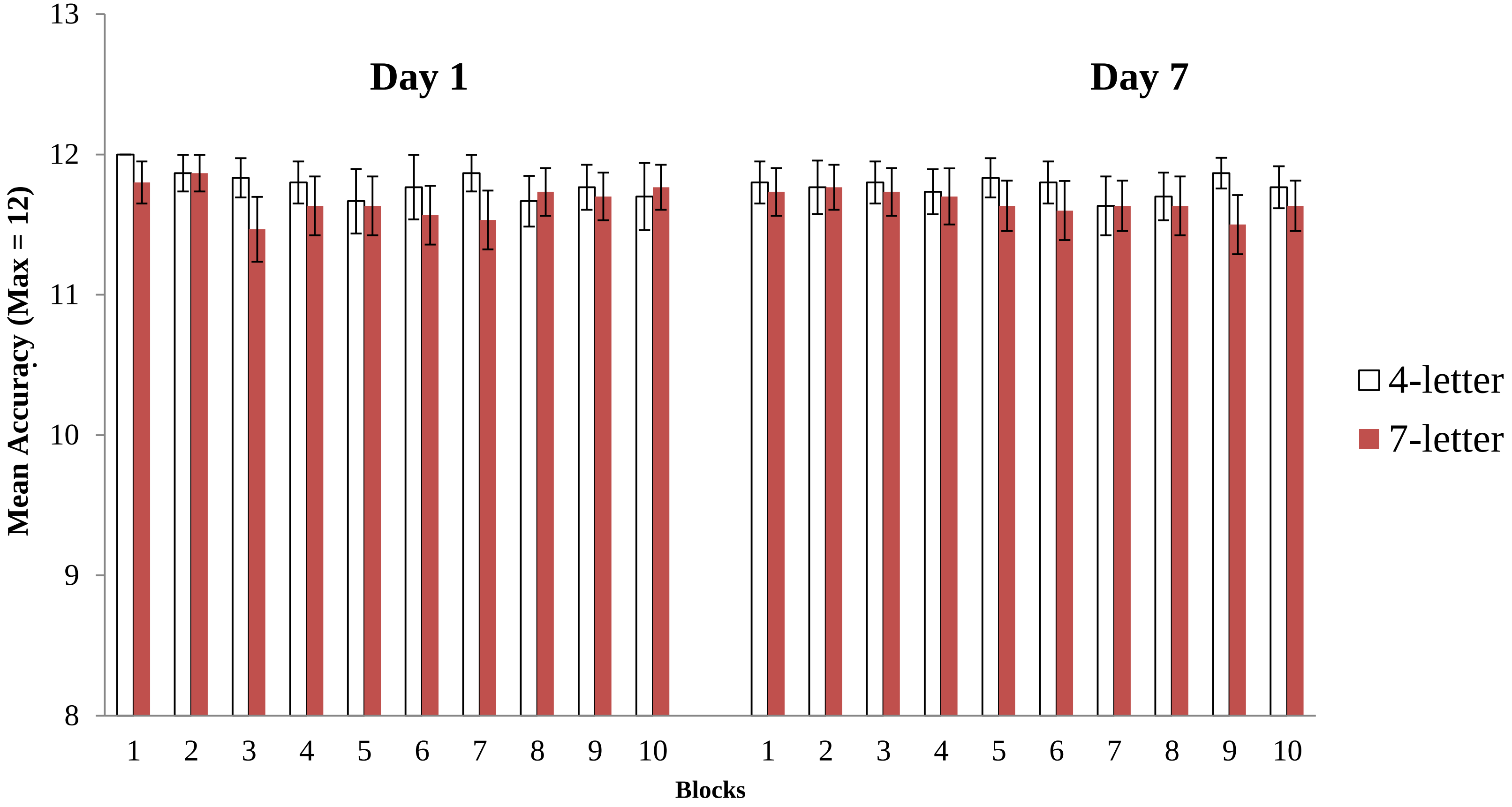

***Figure 5.2*. Naming accuracy to 4- and 7-letter nonwords in dyslexics across two sessions (10 blocks per session). Error bars show 95% CIs.**

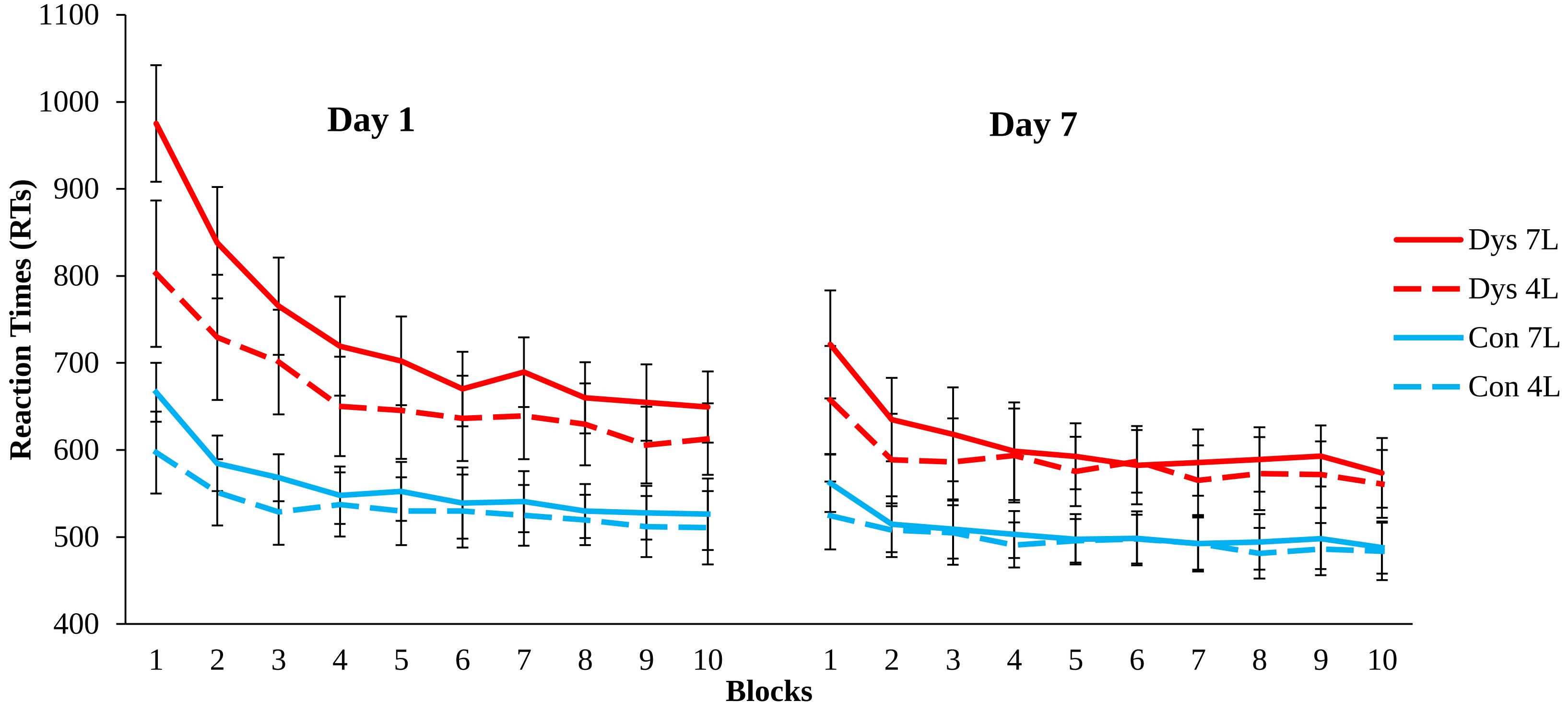

***Figure 5.3.* Naming RTs to 4- and 7-letter nonwords in dyslexics and controls across two sessions (10 blocks per session). Error bars show 95% CIs.**

Table 5.3. Correlations among the predictor variables, and between the predictor variables and naming RTs for 7-letter nonwords in block 1 of day 1.

Variable	1	2	3	4	5	6	7
	Vocab	Literacy	Phon	Wkg mem	Nonverb	Mot	RT
1. Vocabulary	–						
2. Literacy composite	.656**						
3. Phonological awareness	-.403**	-.571**	–				
4. Working memory	.266*	.520**	.432**	–			
5. Nonverbal ability	-.014	-.127	-.175	-.247	–		
6. Motor speed	-.319*	-.452**	-.336**	-.418**	.149	–	
7. Block 1, 7-letter RTs	-.584**	-.739**	-.377**	-.444**	.001	.409**	–

**p* < .05, ***p* < .01. Note that phonological awareness was reverse transformed (thereby reversing the normal direction of correlations). Nonverbal ability and motor speed were square root transformed. RT was log transformed.

In sum, nonword naming RTs in day 1 were slower for the dyslexics than the controls. Both groups showed significant effects of length in the first three blocks, but while the controls showed no difference in naming speed after block 3, the dyslexics continued to show longer RTs to 7- than 4-letter nonwords throughout day 1.

5.3.2.2.3 *Day 7*

The next set of analyses focused on RTs in day 7. As in day 1, there were main effects of Group (faster RTs in the controls than the dyslexics), Blocks (RTs becoming faster across blocks) and Length (faster RTs to 4- than 7-letter nonwords). A significant Blocks x Length interaction reflected an overall reduction in the effect of length across blocks. There were also significant Group x Blocks and Group x Length interactions reflecting more change across blocks and stronger effects of length in the dyslexics than the controls. The 3-way Group x Blocks x Length interaction was marginally significant (p = .06). These interactions were explored further by means of separate analyses of day 7 RTs for controls and dyslexics.

Controls showed effects of Blocks and Length on day 7 with a significant Blocks x Length interaction. Bonferroni-corrected t-tests found a difference in RTs to 4- and 7-letter nonwords in block 1 only. Dyslexics also showed effects of Blocks and Length with a Blocks x Length interaction. In their case, Bonferroni-corrected t-tests found effects of length in blocks 1, 2 and 3, but not from block 4 onwards.

In sum, the controls showed a small effect of length at the start of day 7, but that effect disappeared by block 2. Dyslexics required 3 or 4 presentations in day 7 before they began to show (for the first time) no significant difference between naming RTs to short and long nonwords.

5.3.2.3 Predictors of initial nonword reading speed and novel word learning

The final set of analyses brought together performance on the test battery with two aspects of their naming latency data. Similar to Chapter 3, nonlexical reading skill (decoding) was measured in terms of RTs to 7-letter nonwords in block 1 of day 1 while novel word learning was measured in terms of the change in RTs to 7-letter nonwords from block 1 to block 10 on day 1.

The number of predictor variables was reduced before the regression analyses were run, and some of the variables were transformed to improve the normality of their distributions. There were high correlations among the two word reading tests and the word spelling test (*r*s = .67 - .84, all *p*'s < .001). A composite Literacy score was therefore calculated for each participant by averaging the standardized scores from the WRAT Reading, TOWRE word reading and WRAT Spelling tests. To avoid using nonword reading in one task to predict nonword reading in another task, performance on the TORE-PDE nonword reading task was not included in the composite Literacy score. Substantial correlations were also observed among the four tests of working memory (*r*s = .50 - .56, all *p*'s < .001). A composite Working memory score was therefore computed for each participant by averaging the standardized scores from the four working memory tasks.

Univariate normality was tested for each predictor and the dependent variables (RTs to 7-letter nonwords in blocks 1 and 10 of day 1). Phonological awareness, Nonverbal ability and Motor speed were found to violate the assumption of normality (Kolmogorov-Smirmov test of normality, *p* < .05). Distributions approximated normality most closely when Phonological awareness was reverse transformed (thereby reversing the normal direction of correlations) and Nonverbal ability and Motor speed were square root transformed. RTs were log transformed to reduce skew.

Table 5.3 shows the correlations among the final predictor variables; also the correlations between the predictor variables and RTs to 7-letter nonwords in block 1 of day 1. There were significant correlations among all the predictor variables except Nonverbal ability which did not correlate significantly with any of the other predictors. All of the predictors except Nonverbal ability correlated significantly with RT, with Literacy showing the highest correlation, followed by Vocabulary, Working memory, Motor speed and Phonological awareness. Multicollinearity among the predictor variables was assessed using the variance inflation factor (VIF). VIF scores of less than 4 indicate that the result will not significantly influence the stability of the parameter estimates (Myers, 1990; Olague et al., 2007). VIF scores for the predictor variables ranged between 1.04 and 3.01.

Linear mixed effects modeling was used to explore the ability of Vocabulary, Literacy, Phonological awareness, Working memory, Nonverbal ability and Motor speed to predict initial nonword reading speed and novel word learning. The analyses were conducted in R using the lme4 (Bates et al., 2012) and languageR (Baayen, 2009) packages.

5.3.2.3.1 *Predicting initial nonword reading speed*

The contribution of each predictor variable to predicting RTs for 7-letter nonwords presented in block 1 of day 1 was evaluated by using likelihood ratio tests to compare a model that contained all the fixed and random effects with a sequence of models in which different predictor variables were removed one at a time. These analyses showed that Literacy made a significant independent contribution to predicting nonword naming speed, $\chi^2(10) = 16.12$, $p < .001$; $\beta = -0.005$, $t = -4.30$, $p < .001$. In contrast, Vocabulary, $\chi^2(10) = 2.71$, $p = .096$, Phonological awareness, $\chi^2(10) = 1.41$, $p = .235$, Working memory, $\chi^2(10) = 1.53$, $p = .217$, Nonverbal ability, $\chi^2(10) = 1.37$, $p = .243$, and Motor speed, $\chi^2(10) = 1.12$, $p = .293$, made no independent contributions.

5.3.2.3.2 *Predicting learning*

Novel word learning was assessed in terms of the change in naming RTs for 7-letter nonwords between blocks 1 and 10 of day 1. RTs from both blocks were entered into the analysis. A categorical variable of Time was created to reflect the change in RTs between blocks 1 and 10. A set of predictor variables were then created which were the interactions involving Time with Vocabulary, Literacy, Phonological awareness, Working memory, Nonverbal ability and Motor speed. This makes it possible to evaluate the contribution of each independent variable to predict change in naming RTs to the 7-letter nonwords across blocks (Field, 2012; Shek & Ma, 2011). The effect of the categorical variable of Time was significant, $\chi^2(11) = 516.29$, $p < .001$, reflecting the reduction in RTs from block 1 to block 10. The interactions of Time with Vocabulary, $\chi^2(17) = 6.57$, $p < .05$; $\beta = 0.002$, $t = 2.57$, $p < .05$, and Time with Working memory, $\chi^2(17) = 26.12$, $p < .001$; $\beta = 0.003$, $t = 5.14$, $p < .001$, were also significant. The interactions of Time with Literacy, $\chi^2(17) = 0.71$, $p = .401$, Phonological awareness, $\chi^2(17) = 1.79$, $p = .181$, Nonverbal ability, $\chi^2(17) = 3.65$, $p = .100$, and Motor skill, $\chi^2(17) = 0.10$, $p = .753$, made no independent contributions to predicting RT change across blocks.

In sum, reading latencies for the more difficult, 7-letter nonwords seen for the first time correlated significantly with all of the predictor variables except Nonverbal ability. The highest correlation was with Literacy. When the ability of each of the variables to predict naming RT was assessed in the context of the other variables (in analyses which took into account the differences between participants and items in overall naming speed), only Literacy was significant. Novel word learning was assessed as the change in RTs for 7-letter nonwords between blocks 1 and 10 of day 1. Only Vocabulary and Working memory predicted the degree of learning across blocks in session 1.

5.4 Discussion

The adult dyslexics in the current experiment were all studying at university or in a college of higher education. They performed at a comparable level to typically-reading controls on a test of nonverbal ability (matrix reasoning) but had lower vocabulary scores, slower and less accurate reading and spelling of words, less efficient reading of nonwords, poorer phonological awareness, poorer performance on both verbal and nonverbal tests of span and working memory, and slower motor speed. These findings match other reports in the literature that dyslexics in higher education have cognitive problems that extend beyond reading and writing to wider aspects of linguistic, working memory and motor performance while typically sparing nonverbal reasoning (cf. Bruck, 1992; Callens et al., 2012; Gallagher, Laxon, Armstrong, & Frith, 1996; Hatcher et al., 2002; Smith-Spark, Fisk, Fawcett, & Nicolson, 2003; Smith-Spark & Fisk, 2007; Warmington et al., 2013b). The working memory problems extend to visuospatial as well as verbal tasks (cf. Hachmann et al., 2014; Menghini et al., 2011; Smith-Spark & Fisk, 2007).

The largest difference between dyslexics and controls in the present study (as indicated by the effect size) was on the TOWRE Phonemic Decoding Efficiency test (Torgesen et al., 1999), a test of nonword reading. A great deal of effort is put into teaching phonic decoding skills to dyslexic children in the UK (Rose, 2009). The dyslexics who participated in this study had mastered the letter-sound correspondences of English sufficiently to enable them to read correctly nonwords like *drentcy* and *larquof* on the first encounter, but they were substantially slower than the controls. The results of the TOWRE-PDE indicate that pronouncing unfamiliar nonwords (and, by extension, unfamiliar real words) persists as a problem for dyslexics in higher education (Ben-Dror et al., 1991; Bruck, 1990; Reid et al., 2007; Wolff, 2009).

In the experimental task, the typical readers behaved very similarly to the participants in Experiment 3 of Chapter 2 who were drawn from the same population. Letter length exerted a major effect on reading speeds for nonwords seen for the first time, but the impact of length declined as naming latencies reduced across blocks, becoming nonsignificant from block 4 of day 1 (cf. Maloney et al., 2009). The participants in Experiment 3 of Chapter 2 only required one representation in Day 7 in order to read the nonwords lexically (with no length effect). The same pattern of result was observed in the control group in the current experiment. The current results showed, therefore, that skilled adult readers can create representations of unfamiliar letter sequences after 4 or 5 presentations that allow them to recognize and pronounce the novel 'words' quickly and to process their component letters in parallel.

The dyslexics were substantially slower at reading the nonwords throughout both sessions of the experiment. When the dyslexics read the 7-letter nonwords for the first time in block 1 of day 1, they did so with a mean latency that was over 300 ms slower than the controls. When performance on the 4- and 7-letter nonwords was compared, the dyslexics required 57 ms per letter in order to pronounce a nonword seen for the first time where the controls required just 23 ms per letter (less than half as much as the dyslexics). Ability at reading and spelling real words (‘literacy’) predicted decoding speed across the two groups. When the effect of literacy was taken into account there was no additional effect of vocabulary, phonological awareness or working memory on decoding speed for these particular readers.

The dyslexics in the present study were clearly capable of visual word learning. Figure 5.3 shows that their naming latencies declined across blocks and that their naming latencies to 4- and 7-letter nonwords eventually converged. However, learning occurred considerably more slowly than in the typical readers. Whereas the difference in RTs between shorter and longer nonwords became nonsignificant in the typical readers around the middle of session 1, the dyslexics showed slower naming of longer nonwords throughout session 1, only losing the length effect part-way into session 2 (day 7). The present study confirms, therefore, that the problems with word learning that have been documented in dyslexic children persist into early adulthood, even in high-functioning dyslexics (cf. de Jong & Messbauer, 2011; Ehri & Saltmarsh, 1995; Reitsma, 1983; Mayringer & Wimmer, 2000; Share & Shalev, 2004; Elbro & Jensen, 2005; Thomson & Goswami, 2010).

Importantly, the naming latencies for the dyslexics remained substantially longer than those of the typical readers through to the end of session 2. Figure 5.3 suggests that the difference between the two groups had more or less stabilized by the second half of session 2. Previous studies show that dyslexic university and college students read familiar words aloud more slowly than normal readers (Ben-Dror et al., 1991; Bruck, 1990; Reid et al., 2007; Wolff, 2009): one interpretation of that finding and the present evidence is that no amount of exposure to individual words will allow dyslexic students to reach the point where they can convert them from print to sound as efficiently as typical readers.

In terms of the DRC model of reading (Coltheart et al., 2001), less efficient reading of nonwords in the TOWRE-PDE test and in the experimental task reflects less efficient functioning of the nonlexical route in undergraduate dyslexics than in typical readers. Slower convergence between RTs to shorter and longer nonwords in the dyslexics suggest that the creation of new lexical entries in the orthographic input lexicon and the phonological output lexicon occurs less efficiently in adult dyslexics than typical readers. This results in a slower transition from sublexical to predominantly lexical reading in the dyslexics. Finally, the fact that nonword reading remains slower in the dyslexics than the controls even at the end of session two, combined with the fact that adult dyslexics are slower than controls to read familiar words aloud, indicates that the lexical route also functions less efficiently in adult dyslexics than in typical readers. That could be due to slower operation of the two lexicons or the pathways between them, or it could also be due to less efficient functioning of the final stages involving activating phoneme sequences and converting those sequences into articulation. Problems at the phonological output stage in dyslexics that compromise the functioning of both the lexical and nonlexical routes would be compatible with other evidence for impairments in dyslexics at the speech output stage (see Coltheart, 2005, Havelka et al., 2010, and Ziegler et al., 2008, for discussions of developmental dyslexia within a DRC framework).

Across the two groups, the ability to learn novel words (measured here as the change in RTs to longer nonwords between blocks 1 and 10 of day 1) was predicted by vocabulary and working memory. The finding that vocabulary predicts the ability to learn novel words was consistent with previous studies including Ricketts et al. (2009) and Storkel et al. (2006) and the result of Experiment 2 in Chapter 3. The explanation of how vocabulary can support word learning has already been addressed in the discussion section of Chapter 3.

As regards the contribution of working memory, as noted in the Introduction, studies of children and young adults by Jarrold et al. (2009), Majerus and Boukebza (2013) and Martin and Ellis (2012) found a relationship between working memory and the ability to learn novel words, with working memory apparently related more closely to acquiring new word-forms rather than their meanings. Those observations fit well with the present findings. The DRC model does not engage with the working memory literature directly, but an important part of working memory is the interaction between short- and long-term memory systems exemplified by the interaction between phoneme representations and lexical entries (the phonological output lexicon in the DRC model). Jarrold et al. (2009) and Martin and Ellis (2012) explained the relationship they observed between verbal short-term memory and word learning in terms of individual differences in the ability to maintain accurate phonological representations of novel words. Majerus et al. (2006) argued that maintaining information about the order of phonemes in words is particularly important for successful word learning. In that context, the report by Hachmann et al. (2014) that short-term recall of order information is particularly impaired in dyslexia, may contribute to their word learning problems.

Phonological awareness did not emerge as a predictor of either initial naming RTs or learning when the contributions of the other predictors were taken into account. Research has established that phonological awareness alone is not enough to improve decoding skills: only when phonological training is combined with training on the mappings between letters and phonemes does reading improve (Hatcher et al., 1994; Melby-Lervåg et al., 2012). The fact that phonological awareness was not a significant predictor in nonword reading was not consistent with the findings in Experiment 4 of Chapter 3. This may be due to the fact that this chapter took a step further to include a spelling task in this experiment. Knowledge of the links between letters and sounds may be better captured by the kind of measures of word reading and spelling that went into the Literacy variable in the present study than by phonological awareness based on spoken stimuli and responses.

In conclusion, the results of this experiment show that adult dyslexics in the UK university and further education system continue to experience difficulty reading novel words and nonwords. They are slower to read nonwords aloud than typical readers, requiring more time per letter to pronounce unfamiliar sequences of letters. They show learning of novel words as a result of repeated exposures, but they require more exposures than typical readers before they establish effective lexical representations. Even after multiple presentations their speed of reading aloud is monumentally slower than typical readers. They remain slower than typical readers even at reading familiar words aloud. Across both dyslexic and typical readers, decoding speed for nonwords was predicted by skill at reading and spelling real words ('literacy') while individual differences in word learning were predicted by vocabulary size and working memory. As others have also shown, the problems that adult dyslexics experience extend beyond reading and spelling to word learning, vocabulary, phonological awareness, working memory and even basic motor speed. Taken together, those problems will conspire to make it very challenging for adult dyslexics to function successfully within higher education.

6 Visual word learning in Chinese speakers

6.1 Introduction

Learning to read two languages represents a growing reality for children; almost two thirds of the world's children grow up in bilingual environments (Baker & Jones, 1998; Crystal, 1997). Understanding the cognitive and linguistic processes involved in dual-language reading development has become a central topic in the study of biliteracy acquisition and development (e.g. Cook & Bassetti, 2005; Koda, 2005). Yet, a lot of the work often focuses on using explicit tasks to investigate English word learning in children (e.g. spelling task). The main methodological flaws of these explicit tasks are that they do not tap into the quick and implicit word learning process in second language acquisition (McBride-Chang & Ho, 2005; McBride-Chang, Wagner, Muse, Chow, & Shu, 2005).

This chapter examines the process of visual English word learning in Chinese native adults who speak Cantonese as their first language. This chapter mainly focuses on two questions in second-language acquisition. Firstly, do Chinese native speakers transfer their holistic approach in reading a logographic language (traditional Chinese characters) to reading an alphabetic language (English)? Secondly, given that Chinese character recognition does not implicitly require any phonological awareness (e.g. Siok & Fletcher, 2001) and English word reading in Hong Kong is taught using a 'look and say' method in which teachers pair visual referents, either Chinese characters or simple English words, with their pronunciations. This means that people in Hong Kong are not explicitly taught the instruction of phonological awareness. As a consequence, this may imply that they may have a weaker grapheme-phoneme correspondence (e.g. Holm & Dodd, 1996) . The research question is whether this approach of learning English would guide Chinese readers to read English serially rather than lexically?

6.1.1 Traditional Chinese writing system

This chapter focuses on studying English word learning in people who speak Cantonese as their first language. Cantonese is a monosyllabic and tonal language; the recognition system is composed of two major components, namely the tone recognizer and the base syllable recognizer (Lee, Lo, Ching, & Meng, 2002). Traditional written Chinese is a logographic orthography that deviates greatly from alphabetic writing systems. Unlike many alphabetic languages, a Chinese character is a basic orthographic unit. A character, being monosyllabic, maps onto a morpheme rather than a phoneme in the spoken language. A single character can function as a word and can form a multi-character word with other characters. Characters consist of smaller components (radicals) which may themselves have a pronunciation or meaning.

There are two kinds of radical --- phonetic radicals that suggest the pronunciation of characters and semantic radicals that suggest the meaning of characters (Chen, Fu, Iversen, Smith, & Matthews, 2002; Ho, Ng, & Ng, 2003). Contrary to the alphabetic writing system that is written horizontally running from left to right, Chinese characters are constructed in squares, and the relative positions of the radicals in one character can be top-bottom, left-right etc. Compound characters make up the majority of characters in Chinese, and 85 percent of the compound characters consist of one semantic component (semantic radical, usually appearing on the left side of the character) and one phonetic component (phonetic radical, usually appearing on the right side of the character) (Perfetti & Tan, 1999; Taft & Zhu, 1995). For instance, the character 洋 (/yoeng/, ‘ocean’) contains one semantic radical 氵 that is on the left side of the character, means ‘water’, and the phonetic radical that is on the right side of the character 羊, provides the information of the character’s pronunciation. Fan, Gao, and Ao (1984) suggested that 26.3 percent of semantic-phonetic compounds have an identical pronunciation as its phonetic radical, while other indicate initial, medial, and final tones in the characters in which they appear. Thus, the phonetic radical can only be treated as a reference, and a certain degree of flexibility should be taken into account.

In support of the fact that the phonetic component of Chinese character can only be taken as a reference, Ho and Bryant (1997) conducted a 4-year longitudinal study to examine the relationship between Cantonese Chinese's phonological skills and their success in reading. Initially, 100 Hong Kong Chinese children were tested on visual and phonological skills at the age of 3, before they could read. The findings showed that pre-reading phonological skills significantly predicted the children's reading performance in Chinese 2 and 3 years later, even when the effects of age, IQ, and mother's education were taken into account. They explained the result by suggesting that phonological knowledge helps Cantonese Chinese children to use the phonetic component in Chinese characters.

6.1.2 Second language acquisition

Prior literacy in a first language greatly boosts the ability of a person becoming literate in a second language, regardless of whether the first and second languages are alphabetic or not (Collier, 1989; Nosarti, Mechelli, Green, & Price, 2009; Swain, 1981; Tarone, 1990; Wang, Koda, & Perfetti, 2003). That is, adult language speakers, literate in their first language, can make use of the knowledge and skills of literacy practices from their first language. If one's first language is an alphabetic language, then learning to read a second alphabetic language will increases the phonological associations that can be linked to the same orthographic units (Nosarti et al., 2009). On the contrary, if one's first language is non-alphabetic, then there are other skills that are helpful in learning the second alphabetic language, e.g. phonological skills in syllabic level and visual processing strategies (Wang et al., 2003; Yeung & Chan, 2013). In alphabetic languages, the process of linking phonology to letters relies on phonological awareness. Yet, different orthographies signify their units of phonology differently. Chinese characters represent one-syllable morphemes, not phonemes. As phonological awareness establishes in relation to orthography (Huang & Hanley, 1995), literacy in different orthographies results in differences in phonological awareness.

As mentioned in section 1.8.3 of Chapter 1, the DRC model (Coltheart et al., 2001) of skilled reading in alphabetic orthographies involves the processing of both orthographic and phonological details. Nonetheless, evidence for the use of both phonological and orthographic information in reading is not bound to alphabetic orthographies. There is a general understanding that both activation sources are used to read all orthographies, including alphabetic, syllabic and logographic. For example, both alphabetic and logographic orthographies have been demonstrated to involve phonological coding (Lam, Perfetti, & Bell, 1991). Perfetti and Zhang (1995) tested phonological and semantic interference functions over an interval of brief exposure durations. The time course for semantic interference (in the homophone judgment task) and for phonological interference (in the synonym judgment task) were recorded. It was found that phonological interference resulted with only 90ms stimulus onset asynchrony and semantic interference at 140 ms. This study confirms that though there is no phoneme-grapheme conversion rule in Chinese, phonology is still a component of its identification.

It is well known that those acquiring a second language use the strategies that they found helpful in learning their first language (Coady, 1979; Yeung & Chan, 2013) . The main issue is not whether the skills applied in the first language transfer to the second language, but how these skills are utilized. Difficulties are likely to derive if the skills used in the first language are inappropriate for the second language. For example, readers of a logographic orthography may link a character with the phonological form of a morpheme; hence, they may find it difficult to learn phonological awareness skills that are necessary for reading in an alphabetic language, such as English. As a result, students with a logographic first language may acquire functional literacy using the same strategies they use to read their first language. This leads to the question as to how Chinese speakers acquire English which is the topic that I will now turn to.

6.1.3 How English is taught in Hong Kong

In order to understand how English is taught in Hong Kong, one needs to appreciate how Chinese is taught in the education system. The Chinese literacy programs followed the teaching strategy that rote character learning through ‘look and say’ instruction. The teacher writes a character on the board, indicates the pronunciation, and the children read the character aloud and copy it several times. The teacher provides information about the meaning and use of the character by using it in some sentences.

As with the Cantonese classes for the bilingual children in Hong Kong, the teachers in Hong Kong teach English with the same teaching strategy as teaching Chinese --- by using the ‘look and say’ method. The students do not learn an alphabetic system before they are exposed to English. Students are taught whole-word-to-pronunciation mappings without the mediation of alphabetic decoding (Bialystok, McBride-Chang, & Luk, 2005; Huang & Hanley, 1995; Taft & Chen, 1992). Thus, this instruction relies exclusively on rote memorization of letter sequences.

Yeung, Siegel, and Chan (2013) investigated the effects of a 12-week phonological awareness instruction on 76 Hong Kong young children who were learning English as a second language. The children were assigned randomly to receive the instruction on phonological awareness skills embedded in vocabulary learning activities or comparison instructions which consisted of vocabulary learning and writing tasks but with no direct instruction in the explicit instruction of phonological awareness (just like the ‘look-and-say’ method). The results indicated that children who received the phonological awareness instruction performed significantly better than the comparison group on a wide range of tasks, including spelling and reading (this will be discussed further in section 6.1.4.1.). Interestingly, the regression analyses also showed that phoneme-level phonological awareness significantly predicted the word reading and spelling for the instructional group but syllable-level phonological awareness significantly explained word reading and spelling for the comparison group. This result may implies that though the look-and-say method does not include teaching phoneme-level phonological awareness, readers in Hong Kong may be able to utilize the syllabic-level phonological awareness skills to read nonwords even when they encounter the English novel words for the first time.

As mentioned in section 1.6 of Chapter 1, Breitenstein and Knecht (2002) demonstrated that adults were able to extract rules from words implicitly even though they were not taught the rules explicitly. Using a statistical learning paradigm, Breitenstein and Knecht (2002) showed that adults were able to recognize the pairing of nonwords and pictures implicitly and that performance increased from chance level to

90% correct after 5 days of training and remained good 1 month after training. Thus, there is a possibility that though Hong Kong Cantonese Chinese speakers are not taught the grapheme-phoneme correspondence explicitly, they are able to extract the relationship between grapheme and phoneme on a syllabic level.

6.1.4 Cross-language interaction

The topic of cross-language interaction, or transfer, in word learning has received a good amount of attention over the past few decades. In the psychology domain, this term refers to a statistical correlation between L1 and L2, which implies some communication between the two languages (Wang, Park, & Lee, 2006). Findings on cross-language transfer can be considered by the linguistic interdependence model and the phonological core model.

6.1.4.1 The linguistic interdependence model

The linguistic interdependence model (Cummins, 1979) posits a high level of connection between L1 and L2, in that the L1 skills are actively functioning from the start of L2 learning and therefore provide a foundation for further usage. Thus, the linguistic interdependence model focuses on the similarities between the two languages. This model is supported by Wang, Perfetti, and Liu's (2005) findings. By using four tasks that emphasise the phonological and orthographic processing skills in both Mandarin and English, Wang et al. (2005) found that Chinese children's Mandarin onset matching skill was significantly correlated with English onset and rime matching skills. Pinyin, the alphabetic phonetic system utilized to aid children in learning to read Mandarin characters, was highly correlated with English pseudoword reading. Moreover, Mandarin tone processing, which is non-existent in English, predicted a moderate but significant amount of variance in English pseudoword reading even when English phonemic-level processing skill was taken in account. (Note: Pinyin is not used for teaching Cantonese).

Li, McBride-Chang, Wong, and Shu (2012) found a similar result by using a longitudinal study of 141 Hong Kong Cantonese Chinese children learning to read English as a second language (ESL). The correlation between spelling in Chinese and

English was .64. Longitudinal predictors of English reading comprehension were vocabulary knowledge in both Chinese and English, as well as Cantonese phonological awareness and English word reading. A key point is that in a separate regression, spelling and reading comprehension in L2 (English) were uniquely predicted with spelling and reading comprehension skills in L1 (Cantonese), supporting the notion of transfer for each skill. Given that English spelling is even more lexical than English reading, it is not surprising that the same skills underpin both L1 and L2 spelling.

6.1.4.2 The phonological core view model

In contrast with the linguistic interdependence model, the phonological core view model emphasizes the role of a language-specific phonological core competence in word learning (Geva & Wang, 2001). Bringing this concept to L2 acquisition, this would mean a focus on the L2 phonological system, rather than L1. This model will therefore predict cross-language differences in how phonological representations are related to word reading.

A corresponding pattern of differential effects of phonological awareness in different languages was shown by McBride-Chang, Cheung, Chow, Chow, and Choi (2006) who demonstrated that Hong Kong Chinese ESL children’s Chinese and English vocabulary were predicted by syllable- and phoneme-level awareness, respectively. In the study, the L1s (Cantonese) have the syllables most dominantly characterised in Cantonese Chinese and as a result syllable-level awareness emerged as important. This is contrary to the L2 (English) in which phonemes, not syllables, are the most important predictor in English word learning.

A similar result was obtained by Tong, Tong, and McBride-Chang (2013) in a study in which they investigated the prevalence of Chinese–English language learners who were at risk for reading difficulties in either Chinese or English only, or both, among second and fifth graders in Hong Kong. They examined the metalinguistic skills that distinguished those who were incompetent in reading Chinese from those who were incompetent in reading English. Children who were poor readers of both languages had difficulties in both phonological and morphological awareness. Poor readers of English only were also found to manifest significantly poorer phonological awareness compared to those who were poor readers of Chinese only. This result indicates possible dissociations between the skills that are required for Cantonese Chinese first language word reading and English second language word reading. These findings suggested that the degree to which different metalinguistic skills are essential for reading in different writing systems may depend on the linguistic properties of the particular writing system.

Cheung (1995, 1999) demonstrated the importance of phonological awareness even in English as ESL. These studies found correlations between English phonological awareness and reading in Cantonese Chinese ESL adolescents residing in Hong Kong. Training the Chinese adolescents on English phoneme awareness improved their word learning ability in English. Cheung found in both studies that the phoneme awareness of his Hong Kong adolescent subjects was very limited prior to the intensive training on phonemic analysis, despite their high level of English proficiency. A similar result was obtained by Yeung, Siegel, and Chan (2013) in young children who were residing in Hong Kong. The result indicated that children who received the phonological awareness instruction performed significantly better than the comparison group on English word reading, spelling, phonological awareness at all levels and expressive vocabulary on the posttest when age, general intelligence and the pretest scores were controlled statistically.

Finally, a similar result was reached by McBride-Chang and Treiman (2003). They examined the degree to which young Hong Kong Chinese children used information about letter names and letter sounds to learn English words. Forty children from each of three kindergarten grades (mean age = 3.8, 5.0, 5.9 years old, respectively) were taught to pronounce novel English words that were based on letter-name (e.g, KL = Kale), letter-sound (KL = Kyle), or visual (KL = Bett) cues. By the 2nd year of kindergarten, children performed significantly better in the name condition than the other conditions. The 3rd grade kindergartens also performed better in the sound condition than the visual condition. The results pinpoint the importance of letter-sound knowledge for learning to read English, regardless of native-language or second language acquisition background.

6.1.5 Hong Kong speakers' phonological awareness ability

Holm and Dodd (1996) examined the relationship between first and second language acquisition by identifying the skills developed in the first language that were transferred to the second language. The performance of 10 university students from each of the groups including 1) The People's Republic of China, 2) Hong Kong, 3) Vietnam and 4) Australia were compared on a series of task that assess phonological awareness, reading and spelling skills in English. The results indicated that the Hong Kong students had limited phonological awareness compared to those students with alphabetic first language literacy (including all three groups of students from China, Vietnam and Australia). The reading and spelling tasks showed no differences between the groups on real word reading but the students from Hong Kong had difficulty processing nonwords because of their poor phonological awareness. This result supported the notion that ESL speakers transfer their literacy processing skills from their first language to English. When the phonological awareness required in English had not been developed in the first language, Hong Kong ESL speakers were limited to process English by using an analogy and visual strategy. Therefore, it is expected that Hong Kong ESL speakers will have difficulties with new, or unfamiliar words.

A similar result was obtained by Cheung, Chen, Lai, Wong, and Hills (2001) who compared younger, pre-reading to older, literate children from different linguistic backgrounds on their phonological awareness. Hong Kong (n = 60) and GuangZhou (n = 60) subjects both spoke Cantonese. Guangzhou subjects had early experience with Pinyin (alphabetic) in addition to their logographic Mandarin Chinese reading; the Hong Kong readers read only logographic Chinese. New Zealand subjects (n = 49) spoke English and read the Roman alphabet. The result indicated that 1) the New Zealand pre-readers outperformed their Hong Kong and Guangzhou counterparts on onset, rime, and coda analyses; 2) the Guangzhou children outperformed their Hong Kong counterparts on onset and coda analyses. Finding (2) reflects an effect of alphabeticity in the first learned script as Guangzhou readers had learned alphabetic Pinyin symbols to transcribe the sounds of characters while the Hong Kong readers read logographic characters as their only primary script. Finding (1) appears to be more general than the influence of orthography in that it extends beyond the phonemic level to onset and rime analysis and may relate to subjects' early spoken language experience.

6.1.6 The current experiment and research question

Using the same experimental design of Experiment 3 in Chapter 1, Hong Kong Chinese and British participants read aloud novel English words in two naming sessions that were 7 days apart. Given that Chinese speakers are used to the wholistic approach towards reading Chinese, they may transfer this skill to their second language acquisition (Cummins, 1979; Li et al., 2012; Wang et al., 2005). If that is the case, one would expect to see a similar reduction of length effect that was observed in previous chapters when British speakers read aloud English novel words. In contrast, based on what is observed in Chapter 5, dyslexic students with weaker phonological awareness ability and smaller vocabularies showed 1) slower reaction time, and 2) larger length effects compared to the British control group. Unlike English, Cantonese does not have a phonological alphabetic support system that maps letters to sound. Thus, Cantonese Chinese speakers have always shown a weaker sub-lexical processing ability (e.g. Cheung et al., 2001; Holm & Dodd, 1996). If that is the case, one would expect that the reduction of the length effect in Cantonese Chinese participants would mimic that seen in the dyslexic group in Chapter 5.

6.2 Methods

6.2.1 Participants

Thirty native speakers of English from the control group of Chapter 5, aged 17 – 32 (mean age = 20.70, S.D. = 3.20) and 20 native speakers of Cantonese (Cantonese-English bilinguals, 10 male, 10 female) aged 18 – 26 (mean age = 21.75, S.D. = 2.36) took part in the experiment. All participants were undergraduate or postgraduate students at University of York who were either paid with a small payment or received course credit in return.

Both British and Chinese groups had received a similar number of years of formal education. The mean number of years of education of the English group was 15.95 years (starting with 6-year-old in primary school grade 1, S.D. = 3.22), with a range from 12 to 26 years; where the mean number of years of education in the Chinese group was 15.80 years (S.D. = 2.4), with a range from 12 to 20 years. All students from the Chinese group had been using English as the medium of instruction in all subjects of their primary and secondary school years with the exception of Chinese-related subjects (i.e. Chinese history) and other modern languages (e.g. French and Spanish). All of the Cantonese-English bilinguals had reached an overall score of 7 in IELTS (International English Language Testing System) in order to pursue further education in University of York. None of the Chinese participants have been studying in England for more than 5 years. All participants had normal or correct-to-normal vision with no history of reading problems.

6.2.2 Materials and procedure

Participants attended for two testing sessions, seven days apart. The Materials and Procedure for both sessions were exactly the same as for Experiment 3 in Chapter 2.

6.3 Results

Only RTs for correct responses were analysed. Naming errors, hesitations and failures to activate the voice key accounted for 643 trials (2.79% of the total). RTs shorter than 200 ms or longer than mean plus 2.5 SDs in each block across four- and seven- letter items were regarded as outliers and removed from the analyses of accuracy and RTs. This led to the loss of a further 392 RTs (1.63% of the total), leaving 22965 RTs (95.7% of the total) for analysis. The mean RTs (with standard deviation) in each block on each day for four-and seven-letter nonwords in both Chinese and British group are shown in Table 6.1 along with the percent of final accuracy in each condition. Figure 6.1 shows the naming accuracy of the British group while Figure 6.2 shows the naming accuracy of the Chinese group across the two sessions.

6.3.1 Accuracy

Accuracy was very high (95.1% correct overall and never less than 88.8% correct for either group in any condition or block of trials). Non-parametric Mann-

Whitney *U* test was employed as accuracy was found to violate the assumption of normality (Kolmogorov-Smirmov test of normality, $p < .05$). The Mann-Whitney *U* test found a significant difference between the Chinese and British group on overall accuracy across the two days for 4-letter nonwords, $U(50) = 70$, $Z = -4.58$, $p < .001$. This illustrates the naming accuracy of the 4-letter nonwords across the two days was slightly higher in the British group (mean = 11.76, S.D. = 0.48) than the Chinese group (mean = 11.05, S.D. = 1.19). A significant difference between the Chinese and British group on overall accuracy across the two days for 7-letter nonwords was also observed, $U(50) = 76$, $Z = -4.46$, $p < .001$. This was again due to the 7-letter nonwords across the two days was slightly higher in the British group (mean = 11.72, S.D. = 0.48) than the Chinese group (mean = 11.15, S.D. = 1.13).

When Day 1 and Day 7 accuracies were analysed separately, Mann-Whitney *U* tests found a significant difference between Chinese and British group on overall accuracy across short and long items in Day 1, $U(50) = 24$, $Z = -5.48$, $p < .001$, with the British group having a slightly higher accuracy (mean = 11.71, S.D. = 0.50) than the Chinese group (mean = 11.04, S.D. = 1.14). A significant difference was also observed across short and long items in Day 7, $U(50) = 100$, $Z = -3.96$, $p < .001$, again with the British group having a slightly higher accuracy (mean = 11.77, S.D. = 0.45) than the Chinese group (mean = 11.15, S.D. = 1.18). Despite these differences, accuracy in both groups was very high.

6.3.2 Naming latency

The main analysis focused on the RT data from the experimental task. Figure 6.3 shows the pattern of RTs for correct, trimmed responses across blocks for the Chinese (in red) and the British (in blue). Inspection of Figure 6.3 indicates that naming latencies were slower for the Chinese than the British group throughout the experiment.

In the beginning of the experiment, both groups were slower to read aloud 7- than 4-letter nonwords. The difference in naming RTs for shorter and longer nonwords declined with repetitions, but the Chinese participants seem to have commanded more repetitions to the nonwords before the RTs for shorter and longer items converged.

These indications were explored in a series of ANOVAs. When Mauchly's test of sphericity was significant, the Greenhouse-Geiger correction was applied. Full details of the statistical analyses are presented in the Appendix 5 where effect sizes are reported in terms of the partial eta squared statistic (η^2_p). The important outcomes will be summarized here.

6.3.2.1.1 *Global analysis*

The first ANOVA was a global analysis conducted on the RT data for both testing sessions with Group, Day, Blocks, and Length as factors. There were significant main effects of Group (faster overall RTs for the British than the Chinese group), Day (faster RTs on day 7 than day 1), Blocks (RTs becoming faster across blocks) and Length (faster overall RTs to 4- than 7-letter nonwords). All interactions were significant except the four way interaction between Group, Day, Length and Blocks. Thus there were significant interactions between Group and Length (larger length effects in the Chinese than the British group), Group and Blocks (larger reduction of naming RTs across repetitions in the Chinese than the British group), Day, Group and Length (the reduction of the length effect was quicker in the British compared to the Chinese group in Day 1 than Day 7), Day, Group and Blocks (the reduction of naming RTs through blocks in Day 1 was larger in the Chinese than the British group), and Group, Length and Blocks (the naming RTs of 4- and 7-letter nonwords converged earlier in the British than the Chinese group). These results were examined further by means of separate analyses of RTs in Day 1 and Day 7, including separate analyses of the performance of the British and Chinese groups on each day.

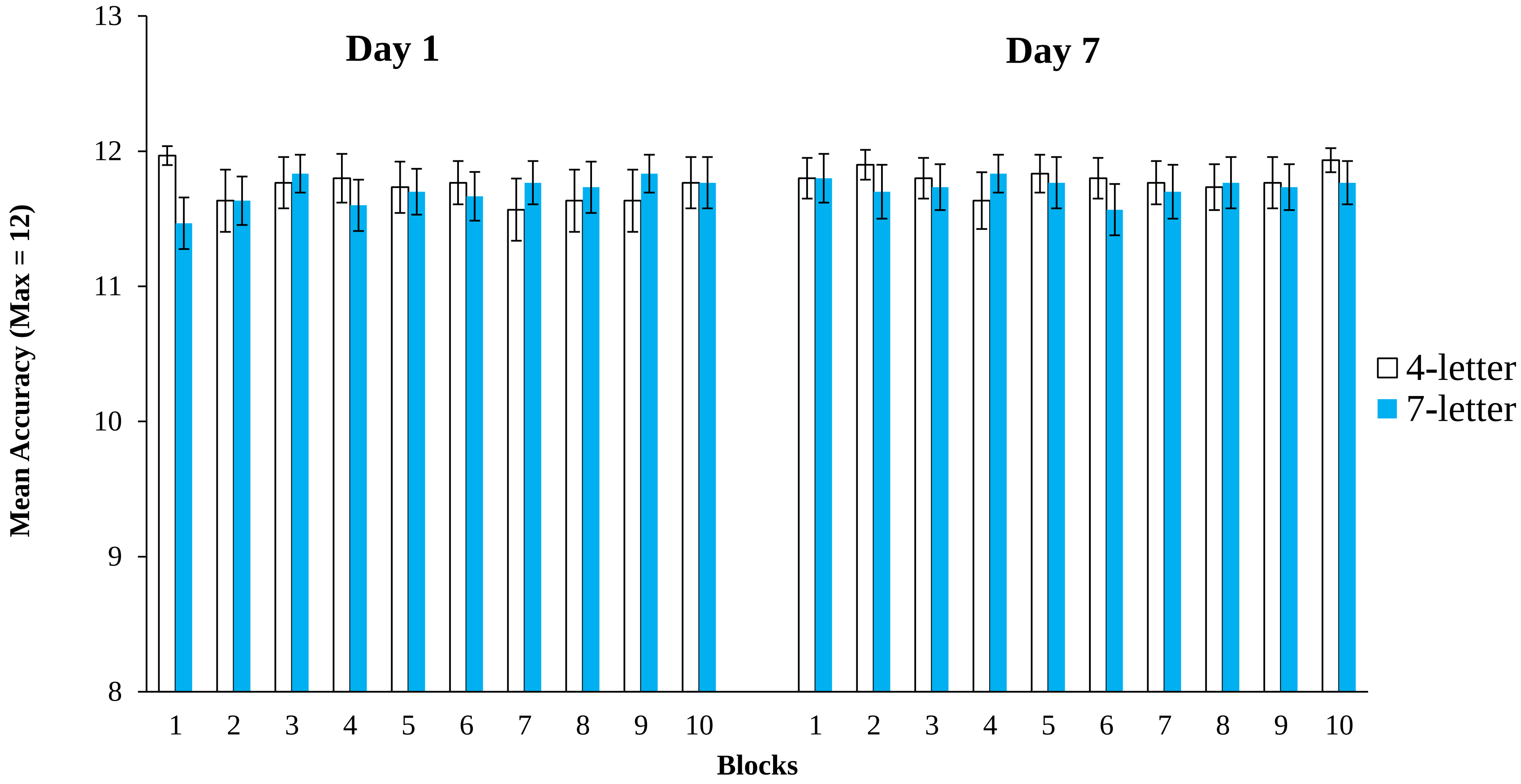

Figure 6.1**. Naming accuracy to 4- and 7-letter nonwords in the British group across two sessions (10 blocks per session). Error bars show 95% CIs.**

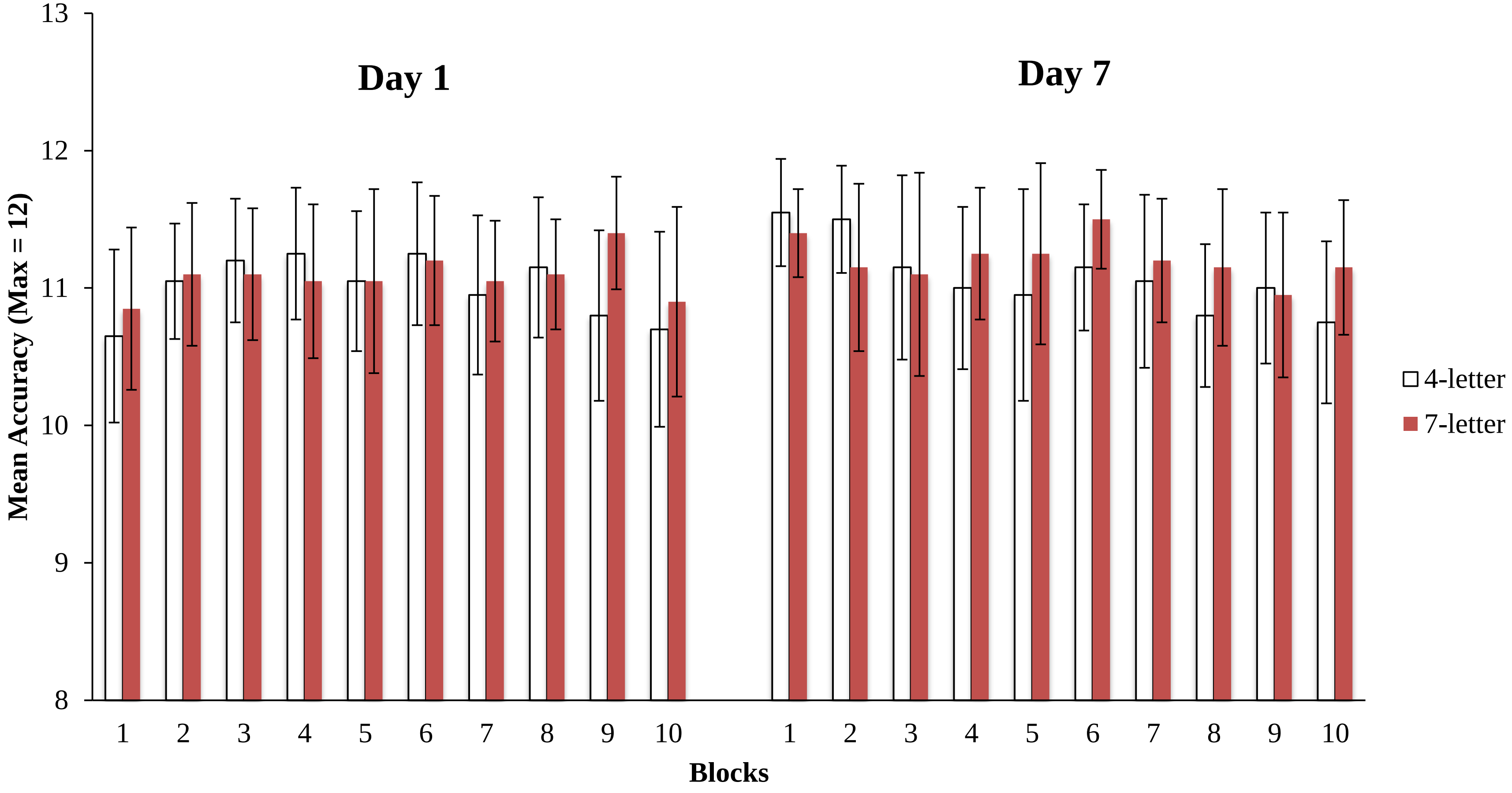

Figure 6.2. **Naming accuracy to 4- and 7-letter nonwords in the Chinese group across two sessions (10 blocks per session). Error bars show 95% CIs.**

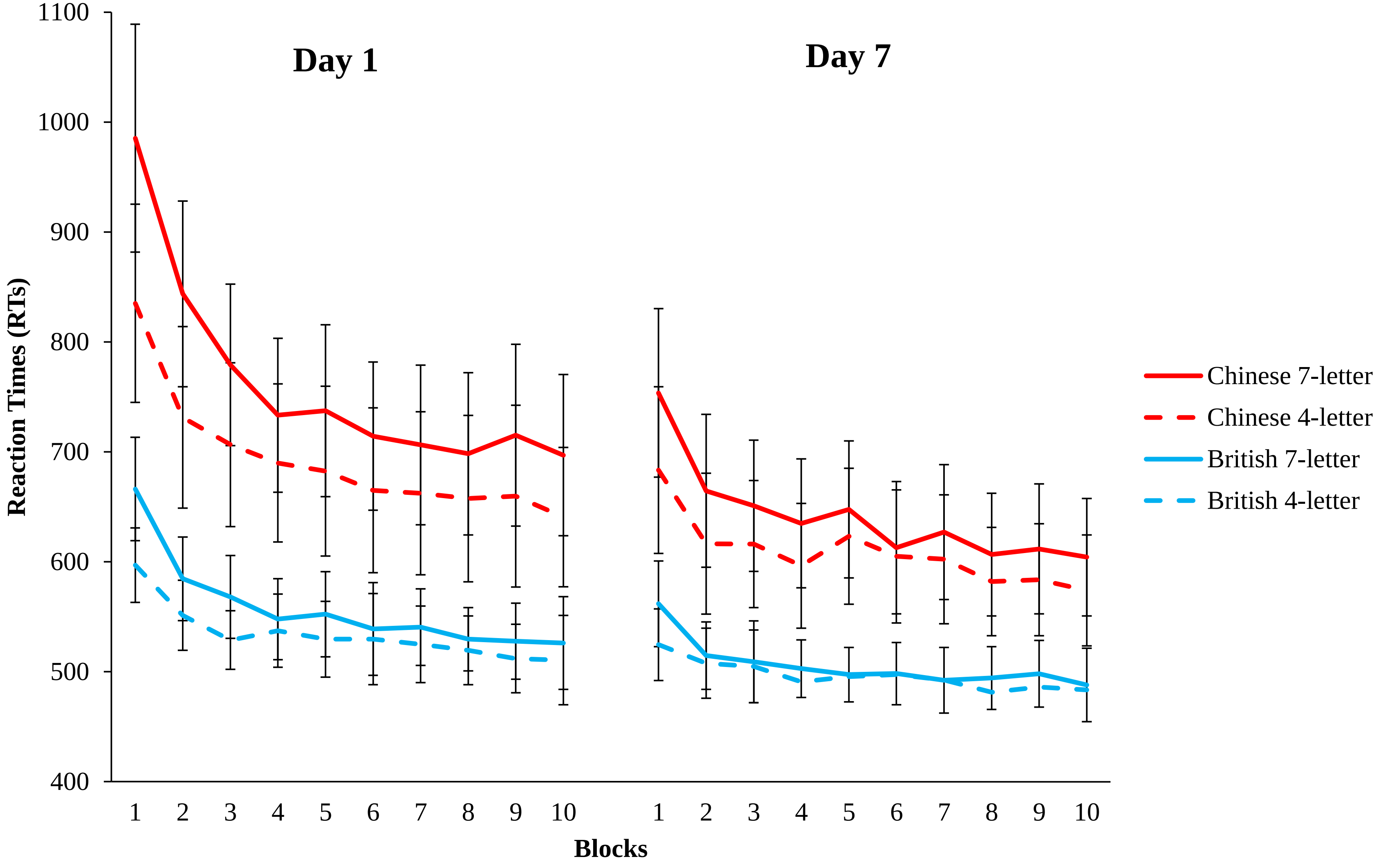

Figure 6.3. **Naming RTs to 4- and 7-letter nonwords in the British and Chinese group across two sessions (10 blocks per session). Error bars show 95% CIs.**

Table 6.1. ***Mean latencies of correct, trimmed responses, standard deviations (S.D.), and per cent correct responses for 4- and 7-letter nonwords in blocks 1 to 10 of day 1 and day 7 in Chinese and British readers.***

	Day 1									
Blocks	1	2	3	4	5	6	7	8	9	10
Chinese										
4-letter nonwords										
Mean RT	835	731	707	690	683	665	662	658	660	641
S.D.	192	176	159	154	165	160	158	162	177	135
Mean Acc.	10.7	11.1	11.2	11.3	11.1	11.3	11.0	11.2	10.8	10.7
S.D.	1.35	0.89	0.95	1.02	1.10	1.12	1.23	1.09	1.32	1.53
% correct	88.8	92.1	93.3	93.8	92.1	93.8	91.3	92.9	90.0	89.2
7-letter nonwords										
Mean RT	985	844	779	733	738	714	706	698	715	697
S.D.	222	180	158	150	167	144	155	158	177	157
Mean Acc.	10.9	11.1	11.1	11.1	11.1	11.2	11.1	11.1	11.4	10.9
S.D.	1.27	1.12	1.02	1.19	1.43	1.01	0.94	0.85	0.88	1.48
% correct	90.4	92.5	92.5	92.1	92.1	93.3	92.1	92.5	95.0	90.8
British										
4-letter nonwords										
Mean RT	597	551	529	537	530	530	525	520	512	511
S.D.	91	85	71	89	92	111	93	84	83	109
Mean Acc.	11.97	11.63	11.77	11.80	11.73	11.77	11.57	11.63	11.63	11.77
S.D.	0.18	0.61	0.50	0.48	0.52	0.43	0.63	0.61	0.61	0.50
% correct	99.7	96.9	98.1	98.3	97.8	98.1	96.4	96.9	96.9	98.1
7-letter nonwords										
Mean RT	666	585	568	548	552	539	541	530	528	526
SD	126	102	101	99	104	113	93	77	93	113
Mean Acc.	11.47	11.63	11.83	11.60	11.70	11.67	11.77	11.73	11.83	11.77
S.D.	0.51	0.49	0.38	0.50	0.47	0.48	0.43	0.52	0.38	0.50
% correct	95.6	96.9	98.6	96.7	97.5	97.2	98.1	97.8	98.6	98.1

	Day 7									
Blocks	1	2	3	4	5	6	7	8	9	10
Chinese										
4-letter nonwords										
Mean RT	683	617	616	597	623	605	602	582	584	574
S.D.	162	137	123	121	132	129	125	105	109	108
Mean Acc.	11.6	11.5	11.2	11.0	11.0	11.2	11.1	10.8	11.0	10.8
S.D.	0.83	0.83	1.42	1.26	1.64	0.99	1.36	1.11	1.17	1.25
% correct	96.3	95.8	92.9	91.7	91.3	92.9	92.1	90.0	91.7	89.6
7-letter nonwords										
Mean RT	754	665	651	635	648	613	627	607	612	604
S.D.	164	149	128	125	133	129	131	119	126	114
Mean Acc.	11.4	11.2	11.1	11.3	11.3	11.5	11.2	11.2	11.0	11.2
S.D.	0.68	1.31	1.59	1.02	1.41	0.76	0.95	1.23	1.28	1.04
% correct	95.0	92.9	92.5	93.8	93.8	95.8	93.3	92.9	91.3	92.9
British										
4-letter nonwords										
Mean RT	525	508	505	491	496	498	492	481	486	484
S.D.	87	85	90	73	79	83	87	85	93	82
Mean Acc.	11.80	11.90	11.80	11.63	11.83	11.80	11.77	11.73	11.77	11.93
S.D.	0.41	0.31	0.41	0.56	0.38	0.41	0.43	0.45	0.50	0.25
% correct	98.3	99.2	98.3	96.9	98.6	98.3	98.1	97.8	98.1	99.4
7-letter nonwords										
Mean RT	562	515	509	503	497	498	492	494	498	488
S.D.	104	82	100	70	66	76	80	77	81	89
Mean Acc.	11.80	11.70	11.73	11.83	11.77	11.57	11.70	11.77	11.73	11.77
S.D.	0.48	0.53	0.45	0.38	0.50	0.50	0.53	0.50	0.45	0.43
% correct	98.3	97.5	97.8	98.6	98.1	96.4	97.5	98.1	97.8	98.1

Note. RT = reaction time (naming latency); Acc. = naming accuracy; S.D. = standard deviation

6.3.2.1.2 *Day 1*

Day 1 RTs were analysed with Groups, Blocks and Length as factors. There were significant main effects of Group (faster RTs in the British than the Chinese), Blocks (RTs becoming faster across blocks) and Length (faster RTs to 4- than 7-letter nonwords). All of the interactions were significant. Day 1 RTs were then analysed separately for 4- and 7-letter nonwords. Both 4- and 7-letter nonwords showed a significant main effect of Group and Blocks along with a Group x Blocks interaction.

In order to understand how the effects of Length and Blocks differ in the British and Chinese group, the Day 1 RTs were analysed separately for the British and Chinese groups. The British group showed significant main effects of Blocks and Length with a Blocks x Length interation. Bonferroni-corrected t-tests were used to compare RTs to 4- and 7-letter nonwords in blocks 1 to 10. The effect of length was significant for the British group in blocks 1, 2, and 3 (all reached $p < .001$) but was no longer significant from block 4 onwards. The Chinese group also showed effects of Blocks and Length along with a Blocks x Length interaction. In their case, Bonferroni-corrected t-tests found effects of length in blocks 1-3, 5, and 7-10 (all reached $p < .005$), with marginally significant effects in blocks 4 and 6 (see Appendix 5).

In sum, nonword naming RTs in day 1 were slower for the Chinese than the British group. Both groups showed significant effects of length in the first three blocks. Nevertheless, while the British group showed no difference in naming speed after block 3, the Chinese groups continued to demonstrate longer RTs to 7- than 4-letter nonwords all through day 1.

6.3.2.1.3 *Day 7*

The next set of analyses focused on RTs in day 7. There were main effects of Group (faster RTs in the British than the Chinese group), Blocks (RTs becoming faster across blocks) and Length (faster RTs to 4- than 7-letter nonwords). A significant Blocks x Length interaction reflected an overall reduction in the effect of length across blocks. There were also significant Group x Blocks and Group x Length interactions reflecting more change across blocks and larger of length in the Chinese than the British group. The 3-way Group x Blocks x Length interaction was not significant.

Though the three-way interaction between Group, Blocks and Length was not significant, Day 7 RTs were analysed separately for 4- and 7-letter nonwords in order to understand the significant interaction between Length and Groups. Both 4- and 7-letter nonwords showed a significant main effect of Blocks and Group along with a Blocks x Group interaction. In order to understand how the effect of Length and Blocks differ in the British and Chinese group, the Day 7 RTs were then analysed separately for British and Chinese group.

British group showed effects of Blocks and Length on day 7 with a significant interaction between Blocks and Length. Bonferroni-corrected t-tests found a difference in RTs to 4- and 7-letter in block 1 only. The Chinese group also showed effects of Blocks and Length with a Blocks x Length interaction. In their case, Bonferroni-corrected t-test found effects of length in blocks 1 to 5 (all reached $p < .005$), but not from block 6 onwards.

In sum, the British group showed a small effect of length at the start of day 7, but that effect disappeared by block 2. The Chinese group required 5 presentations in day 7 before they began to show (for the first time) no significant difference between naming RTs to short and long nonwords.

6.4 Discussion

The Cantonese native speakers in the current experiment were all studying at university in England. Their naming latency pattern was just below the typically-reading controls. The Cantonese Chinese readers demonstrated a very high accuracy while reading the nonwords aloud. The Chinese speakers in the present study were clearly capable of visual word learning. Figure 6.3 illustrates that their naming latencies declined across blocks and their naming latencies to 4- and 7-letter nonwords eventually converged. Yet, learning occurred noticeably more slowly than in the British readers. Whereas the difference in RTs between shorter and longer nonwords became nonsignificant in the British readers around the middle of session 1, the Chinese readers showed slower naming of longer nonwords through session 1, only losing the length effect in the middle of session 2 (day 7). The present study confirms, therefore, that the problems with learning new English written words inefficiently in Cantonese Chinese speakers persist to early adulthood, even in highly literate adults (c.f. Cheung, 1995, 1999; Cheung et al., 2001; Holm & Dodd, 1996; McBride-Chang et al., 2006).

Cantonese Chinese speakers were substantially slower at reading the nonwords throughout both sessions of the experiment (η_p^2 = .362). When one reads nonwords for the first time, these nonwords must be pronounced using relatively serial letter-sound conversion processes. There is no alternative way to read the nonwords for the first time. When the Chinese speakers read the 7-letter nonwords for the first time in block 1 of day 1, they did so with a mean latency that was over 300 ms slower than the British native speakers. When performance on the 4- and 7-letter nonwords was compared, the Chinese speakers required 50 ms per letter in order to pronounce a nonword seen for the first time whereas the controls required just 23 ms per letter (less than half as much as the Chinese speakers).

Importantly, the naming latencies for the Chinese speakers remained substantially longer than those of the British readers through to the end of session 2. Figure 6.3 suggests that the difference between the two groups had stabilized by the second half of session 2. In line with previous studies, the Cantonese Chinese native speakers in the current experiment are slow inefficient letter-sound (phonics) readers (Holm & Dodd, 1996; Cheung et al; 2001). This may be a reflection of how they were taught to read English and their poor phonological awareness in English (as this experiment did not test measure the phonological awareness skills directly, this can only be a logical suggestion based on the literature). The Chinese speakers are also slower to create lexical representations and read the nonwords as whole units. Based on the findings of Chapter 5, this problem may be linked to smaller vocabulary size rather than to the skills of phonological awareness.

6.4.1 The interdependence model in English word learning

Given that Chinese is a logographic language, we would expect to see a reduction in the length effect if the Chinese speakers transfer the skills that they acquired in reading Chinese to English word learning. Thus, the result of the Cantonese Chinese speakers did *not* provide evidence that the skills developed and strategies used in L1 are transferred to L2 (Coady, 1979) in this particular task that they had to read novel words aloud. This is not to say that there are no transferable skills that can be transferred from Chinese to English. In fact, Li et al. (2012) found Chinese vocabulary knowledge was a longitudinal predictor of English reading comprehension. Thus, there are transferable skills between Chinese and English languages but possible only when meaning is involved.

This experiment also did not investigate the effects of length on naming latency for very familiar (high-frequency) English words of Chinese native speakers. If Chinese readers also show larger length effects for high-frequency words than English native speakers, this will imply that the Chinese readers are not reading English words holistically.

6.4.2 The phonological core view model in English word learning

In contrast with the interdependence model, the phonological core view model focuses on the specific linguistic skills that are important distinctively to L1 and L2 (Geva & Wang, 2001). Bringing this concept to this experiment, this would mean an emphasis on the skills that are important in L2 (English) acquisition, which is phonological awareness in English nonword reading. The Cantonese Chinese participants learn English from a ‘look-and-say’ method (Bialystok et al, 2005) since they were 6 years old. They did not have the opportunity to acquire phonemic awareness before they learned English and the ‘look-and-say’ learning instruction does not facilitate phonemic awareness at that stage either. Given that Cantonese speakers read English nonwords much slower than the British native speakers when they first saw the English nonwords, this demonstrates that with little explicit instruction on phonological awareness, at either the subsyllabic or phonemic levels, this makes it difficult for them to assemble phonology to allow them to process nonwords in an efficient way. This result is in line with the previous literature that there are dissociations between the skills that are required for Chinese and English word learning (Tong et al., 2013; McBride-Chang et al., 2006) .

The performance of the Cantonese Chinese participants in the current experiment was similar to that of the dyslexic adults in Chapter 5. Both groups were equally slow in reading nonwords and only managed to start reading the nonwords in a holistic way for the first time in session 2 (day 7). This is not to say that Hong Kong speakers are dyslexic, as the Hong Kong subjects do not show any apparent developmental or acquired deficit. The Cantonese speakers probably would be able to acquire phonological awareness skills if they had been taught an alphabetic system in a systematic way. They have a phonological awareness deficit only due to the fact that they have not been exposed to the phonological segmentation teaching and their experience in Chinese does not encourage the development of phonemic (rather than syllabic) awareness. This result pinpoints the importance of letter-sound knowledge for learning to read English, regardless of native-language or second language acquisition (McBride-Chang & Treiman, 2003; Cheung, 1995; 1999; Yeung et al, 2013).

An important note is that the Hong Kong readers in the current experiment show an adequate level of letter-sound conversion skills that would allow them to read the nonwords aloud even they had not seen the novel words before. This may be due to two reasons. Firstly, the ‘look-and-say’ teaching method may have facilitated syllabic-level phonological awareness. Yeung et al. (2013) found that phoneme-level phonological awareness significantly predicted the word reading and spelling for the group that received explicit instruction on phonological awareness while syllable-level phonological awareness significantly explained word reading and spelling for the control group that received ‘look-and-say’ instruction. Thus, there is a chance that the ‘look-and-say’ teaching instruction aids Chinese speakers to read the nonwords on the syllabic level. Furthermore, though Cantonese Chinese native speakers did not receive any explicit instruction on letter-sound knowledge, they might have acquired some fundamental of grapheme-phoneme correspondence implicitly. This is relevant to Breitenstein and Knecht’s (2002) findings that adults were able to extract rules from novels words implicitly.

Secondly, this may relate to the role of status in L2 English learning. All the participants in the current experiment were international students who were originally from Hong Kong and were studying in the University of York. Most of these participants would come from families that have a higher income and education levels and have more resources to facilitate English language and literacy learning. For instance, it is not uncommon in Hong Kong for families from middle and upper socioeconomic status to hire domestic helpers, to live with them and take care of several household chores including cooking, tidying and child care. Many of these domestic helpers are from the Philippines and Malaysia and they can communicate in English fluently. All participants in the Chinese group had at least one domestic helper since birth and the domestic helpers would communicate with them in English. Furthermore, as all the participants from the Chinese group came from a middle-class status, their families can afford to hire tutors to facilitate their English learning. These factors may contribute to the adequate level of phonological awareness in the Hong Kong participants.

The current experiment, combined with the literature revised in the introduction, have the following practical implications. Firstly, the relatively weak nonword reading ability in Cantonese native speakers that was observed in previous studies persists to adulthood (Cheung et al, 2001; Yeung et al, 2013, Cheung, 1995, 1999). This means that Chinese speakers may have difficulty in studying a new subject that has lots of vocabularies in English, e.g. neuroscience. This is due to the fact that it will take them much longer to read a page of textbook compared to typical British native speakers. Secondly, while phonological awareness is not a prerequisite for the achievement of high levels of literacy in English, training in phonological awareness allows the use of phonics for learning new words efficiently. Given that Cantonese Chinese speakers do not have the opportunity to learn phonological awareness in their own education system; it may be beneficial to include explicit phonological awareness and vocabulary instruction in any pre-sessional university education courses. Thirdly, while the rich oral language environment may help to develop a very fundamental understanding towards syllabic segmentation, a more sophisticated level of phonological awareness requires explicit instruction for development.

One of the issues remaining to be resolved is that the cognitive profile of the participants was not recorded in the current experiment. Thus, is it not certain what cognitive skills Hong Kong readers utilize to help them to learn English (e.g. working memory). Furthermore, future experiments should include different tests that tap into phonological awareness in different levels, including the simple phoneme deletion task and the more difficult ones (e.g. spoonerism). This may show that given that Cantonese Chinese speakers have a rich environment to learn English in Hong Kong, this may aid their phonological awareness development in a very fundamental but not a sophisticated level. Thus, one would expect that Cantonese Chinese speakers may be able to complete the basic phoneme deletion task but not the spoonerism task. Furthermore, based on the result of Chapter 5, vocabulary was the important predictor of English word learning. Chinese students presumably have a large vocabulary size in Chinese, but a smaller English vocabulary than English native speakers. This suggests that vocabulary size in English is what predicts English word learning. This can be tested in a future experiment by measuring L1 and L2 vocabularies separately. Nonetheless, the paradigm developed here can be considered as a tool to understand the process of visual word learning in Cantonese native speakers and to understand how their native language interacts with their second language acquisition.

In conclusion, the current experiment shows the Cantonese Chinese native speakers who study in the UK university continue to experience difficulty reading novel words and nonwords. Not only are they are slower to read nonwords aloud than British native speakers, but they also require more time per letter to read the unfamiliar sequences of letters aloud. Chinese native speakers do show learning of novel words to a certain extent as a result of repeated exposures, but they required triple the amount of exposure compared to the British native speakers before they can process the long novel words in a holistic way for the first time. Even after 20 presentations of the novel words, the Chinese native speakers still showed a much slower naming speed compared to the

British native speakers. This result is attributed to the fact that Cantonese Chinese speakers do not acquire phonological awareness knowledge in both their L1 and L2, thus, these skills are simply unavailable to transfer from L1 to L2. This current result along with the previous literature pinpoint the importance of explicit instruction of phonological awareness for learning to read English, regardless of native-language or second language acquisition (McBride-Chang & Treiman, 2003). Despite being highly literate, the fact that Cantonese native speakers do not receive any explicit phonological awareness training may imply that they may demonstrate similar learning difficulties as with students with dyslexia (e.g. spelling problems), which may pose a problem when they have to learn a subject that involves lots of vocabularies in higher education.

7 General Discussion

7.1 Summary of key findings

The research reported in this thesis aimed to investigate the process of orthographic and phonological word learning in adults. The aim of the thesis was addressed in 9 experiments over 5 chapters. First, Chapter 2 (Experiment 1, 2, 3) developed the fundamental learning paradigm that used speed of reading aloud as the main measure, especially the reduction in naming RTs to short and long novel words through repetition and the convergence of RTs to short and long items. Chapter 3 (Experiment 4, 5, 6) then utilized the same learning paradigm to understand the role of phonology in visual word learning, with a view to ascertaining whether reading aloud training would be found to be more efficient than hear-and-repeat training with and without distractors. Chapter 4 (Experiment 7) then involved the repeated presentation of interleaved high-frequency words, low-frequency words and nonwords to native speakers of English in two testing sessions 28 days apart. The theoretical purpose was to understand the relative effects of length on naming latencies for high-frequency words, low-frequency words and nonwords, the extent to which those latencies (RTs) converge for shorter and longer words and nonwords, and the persistence of training/repetition effects over a 28-day retention interval. Chapter 5 (Experiment 8) brought these theories in a more applied context to understand orthographic word learning in adults with dyslexia who are in higher education. Finally, Chapter 6 (Experiment 9) examined the process of visual English word learning in Chinese native adults who speak Cantonese as their first language in order to understand whether Chinese native speakers transfer their holistic approach in reading a logographic language (traditional Chinese characters) to reading an alphabetic language (English).

2.6.1 Chapter 2: Visual word learning in skilled readers of English

The three experiments in Chapter 1 showed the same pattern of results for novel words read aloud 10 times in 10 separate blocks within a single session. In the first block of trials, when the nonwords were read for the very first time, naming RTs were slow and the naming RTs difference between 4- and 7-letter nonwords was large. This is in line with the literature that there is often a substantial length effect in naming RTs of nonwords (Juphard et al., 2004; Mason, 1978; Valdois et al., 2006; Weekes, 1997; Ziegler et al., 2001)

In all three experiments of Chapter 2, naming RTs declined with repetition of the trained nonwords across blocks. The reduction was larger for the long items (7-letter than 4-letter nonwords). As a result, the RTs to short and long nonwords converged across repetitions. The pattern for the first four blocks was similar to those reported in Maloney et al. (2009) for nonwords of 3 to 6 letters. Experiment 1 of Chapter 1 extended Maloney et al. (2009) experiment by extending the training session beyond four presentations to 10. Mean RTs asymptoted at around block 6.

Experiment 2 of Chapter 2 then explored how the blocking effect affects the pattern of results that was observed in Experiment 1. In experiment 2, trained nonwords were interleaved with untrained nonwords in block 1 and 10. Though there was some general improvement for the naming of untrained nonwords in block 10, the gradual reduction of length effect could only be observed in the nonwords that were trained for 10 blocks. The fact that untrained nonwords were read more quickly when they were interleaved with trained nonwords in block 10 than when they appeared in block 1 with nonwords that were being read for the first time in block 1 could be explained by the blocking effect. Lupker et al. (1997), Rastle et al. (2003), Taylor and Lupker (2001), and Reynolds et al. (2012) explained that participants set a criterion for the speed of responding to stimuli in a block based on the combination of easy and difficult items within the block. When the items are all easy, the criterion will be relatively shorter and RTs consequently would be faster, and vice versa.

Experiment 3 of Chapter 1 then aimed to understand the retention of learned materials by asking participants to return for a second training session that was 7 days after the first learning session. Naming RTs increased between the end of the first testing session and the start of the second session. But after two or three presentations in that second session, RTs had decreased back to the level observed at the end of the first session and the convergence of RTs to short and long nonwords was achieved. This result was consistent with the result that was obtained by Salasoo et al. (1985) who found a warm-up effect in which participants showed lower accuracy of learned pseudowords before they re-familiarized themselves with the learned letter strings again a year after they learned the new items. Taken together, these results suggest that though the absence of exposure to the novel words may lead to a small amount of decay of representations between the end of one session and the start of the next, a few presentations allow the representations to strengthen further.

The reduction of the length effect through repetitions can be explained by the DRC model (Coltheart et al., 2001). When the participants first saw the nonwords in block 1 of the first training session, they can only read these nonwords based on the grapheme-phoneme correspondence (the non-lexical route). As the nonwords are processed in the sequential, left-to-right form, the robust length effect was the inevitable result by processing from the non-lexical route. Moving forward to blocks 2 to 7 in the first training session, though participants would still process the nonwords in a serial way, they also started creating lexical entries in the orthographic input and phonological output lexicons. As the lexical route processes relatively slowly, a small but significant contribution from the non-lexical route was still observed. During blocks 8 to 10 of the first training session, the naming performance was fully dominated by the lexical route. Though the non-lexical route cannot be switched off, given that the lexical route operates very quickly, verbal response is delivered by the lexical route before the non-lexical route is able to produce any response. Thus, no significant length effect was observed between blocks 8 to 10 in the first training session. Experiment 3 of Chapter 2 showed that once participants have built lexical entries in the orthographic and phonological lexicons, the representations can lasts for 7 days without further exposure to the learned materials.

2.6.2 Chapter 3: The role of phonology of visual word learning

The three experiments in Chapter 3 utilized the learning paradigm that was introduced in Chapter 2 to investigate orthographic learning in skilled adult readers, integrating theories of reading development with the skilled reading literature. The 'item-based' account of lexical acquisition put forward by theorists including Share (1995, 1999, 2004), Ehri (1989, 1992) and Perfetti (1992, Perfetti & Hart, 2001) was expanded to address the role of phonology in orthographic learning. Share's theory of phonological recoding (print-to-sound correspondence) as a lifelong self-teaching mechanism is extended to explore the potential role of feedback from phonology in the process of orthographic acquisition of new words.

Experiment 4 was an extension of Maloney et al. (2009) study. All participants read aloud the novel words in Block 1, there were two types of training after Block 1 and participants had to read aloud novel words again in the final block. There were two types of training, hear-and-repeat and read aloud, all participants received both training with half the participants going through read aloud training first before they received the hear-and-repeat training. In both orthographic and phonological training conditions, RTs became shorter across blocks as the nonwords became familiar. Same as Experiment 1, 2, and 3, this result was more apparent for the 7-letter nonwords. Learning was better in the orthographic training conditions in which participants had the benefit of receiving training on both the orthography and the phonology of the stimuli but training the phonology was sufficient to build representations in the mental lexicon.

Linear mixed effects found that literacy and phonological awareness made a significant contribution to predicting 7-letter nonwords naming speed in block 1 even when vocabulary and rapid digit naming were taken into account. This result is consistent with the developmental (e.g. Bowyer-Crane et al, 2008; Muter et al., 2004; Ricketts et al., 2009) and adult (Young et al., 2002) studies showing that phonological skills are still a crucial factor that affect the speed of reading nonwords when children proceeds to their adulthood. As reported by Braze et al. (2007), and Nation and Snowling (2004), expressive vocabulary is a significant predictor of the change in naming RTs for 7-letter nonwords between block 1 and 10 of day 1 in both orthographic and phonological training conditions even when literacy, phonological awareness were taken into account. The result extends Ouellette's (2006) finding that vocabulary depth/expressive vocabulary is not only related to reading comprehension but also affect how well participants can build representations in the mental lexicon.

Experiment 5 then aimed to demonstrate that the learning effect that was observed in Experiment 4 was not due to the cross-task correlation (Lovett et al., 2000). Thus, Experiment 5 adopted a between-subject design in which participants are trained either in the orthographic or the phonological learning condition. Result showed that RTs to trained nonwords in both the orthographic and phonological training conditions showed a similar pattern to that observed in Experiment 4, with both conditions showing a larger reduction of length effect for the trained compared to the filler items. Training on the phonology of novel words had successfully built representations in the mental lexicon; yet, orthographic training had yielded greater improvement compared to the phonological training condition. There were also significant differences between the hear-and-repeat and read aloud conditions. While all 4- and 7-letter trained and fillers items showed significant improvement in the read-aloud condition, only the trained 7-letter items were marginally significant for the hear-and-repeat condition. This may imply that the blocking effect that was mentioned in Experiment 2 of Chapter 2 was specific to orthographic training.

Experiment 6 aimed to alleviate orthographic activation during phonological learning by including orthographic (letter strings) and non-orthographic (pictures) distractors in the hear-and-repeat condition. There were four conditions in this experiment, which the read aloud and hear-and-repeat conditions were the same as those in Experiment 4 and 5, Experiment 6 included two additional conditions, namely 1) Hear-and-repeat with orthographic distractors in which the 4-letter strings would change every 500 ms on the screen and 2) Hear-and-repeat with non-orthographic distractors in which facial pictures would change every 500ms on the screen. In order to assure that participants look at the distractors (rather than ignoring them), a red dot was used as catch item and would appear on the screen in 6% of the experimental session. Participants pressed a button of a response box when it appeared. The first block of reading aloud training in Experiment 4 and 5 had also been eliminated in Experiment 6 in order to reduce the possibility that participants use the spelling of the auditory novel words as a strategy to learn them. Only 7-letter items were included in Experiment 6.

As with the result of Experiment 5, Experiment 6 showed that naming RTs to trained items were faster compared to untrained items. All four conditions found significant differences between trained and untrained items, with RTs showing faster RTs to trained than untrained items. Focusing on the difference of the naming latency of the trained and untrained items, there was a significant difference between the read aloud and hear-and-repeat with orthographic distractors conditions and the read aloud and hear-and-repeat with non-orthographic distractors conditions. High accuracy rate of the red dot demonstrated that participants were indeed paying attention to the distractors on the screen. In accordance to the ‘lexical quality hypothesis’ (Perfetti, 1992), Experiments 4, 5, and 6 showed that learning was better in the orthographic training conditions which participants had the benefit of receiving training on both the orthography and the phonology of the stimuli. Yet, training the phonology of the new words is sufficient to build representations in adults’ mental lexicons. This is in line with previous developmental (Reitsma, 1983) and adult literature (Chalmers and Burt, 2008; Sandak et al., 2004) that as phonology is an essential part of learning new words, training the phonology of new words is sufficient to help participants to build representations in the mental lexicon.

2.6.3 Chapter 4: Reading and lexicalisation in English

Chapter 4 utilized the same learning paradigm that was developed in Chapter 2 to understand how the newly learned items integrate with existing knowledge in the mental lexicon with high- and low-frequency words. Unlike Chapter 2 and 3, high-frequency and low-frequency words were integrated with nonwords in Experiment 7 and the two sessions are 28 days apart instead in order to understand whether the newly learned words showed good retention over an interval of 4 weeks. Result showed naming RTs were faster to high- than low-frequency words and faster to words than nonwords. Those differences were again larger for longer items and diminished across blocks and days as items were repeated.

The result of Experiment 7 could be explained by the DRC model (Coltheart et al., 2001). Similar to the result of Experiment 1, 2, and 3 of Chapter 2, naming RTs of nonwords diminished across the first 6 blocks on day 1 before reaching asymptote. This implies it took around 6 blocks for participants to build representations in the orthographic input and phonological output lexicons. This result is in line with the literature that children (Ehri & Saltmarsh, 1995; Share, 1999) and adults (Gaskell & Dumay, 2003; Maloney et al., 2009) showed rapid word learning. The process of how nonwords are processed through the nonlexical route had already been mentioned in section 7.1.1. Thus, this section focuses on the result of the high- and low-frequency words. Though low-frequency words had already build representations in the orthographic input lexicon and the phonological output lexicon, given that lexical conversion of low-frequency words from print to sound is slow and may overlap in time with the delivery of the rule-based pronunciation by the nonlexical route, a marginal effect of length was observed when participants read the low-frequency words for the first time. For the result of the high-frequency words, though the nonlexical route cannot be switched off, the lexical conversion of high-frequency words operates so quickly that a response was made before the nonlexical route delivers a rule-based pronunciation. This explains why a length effect was not found for the high-frequency words for the Experiment 7.

The fact that RTs of both high- and low-frequency words also improved through blocks in day 1 can be explained by repetition priming effect. There have been many reports suggesting that recognition of familiar English words benefit by repeated presentation (i.e., repetition priming under conditions where each presentation of a word is obviously visible rather than masked). Stark and McClelland (2000) mixed words, pseudowords (word-like nonwords) and nonwords (letter strings that looked unlike real words) in the CID task, they had also found words being identified faster than pseudowords and pseudowords were recognized faster than nonwords; they also found a significant main effect of repetitions and an interaction between stimulus type and repetition. The result of the current experiment is in line with the literature that the effect of repetition priming is quite long-lasting (Wiggs et al., 2006).

Experiment 7 showed some increase in RTs at the start of day 28 for the English nonwords but the benefits of repetition were reinstated in block 2 of day 28. It only took participants one block to read as quickly as in block 10 in day 1. Even with the increase for nonwords across the delay, RTs at the start of day 28 were faster than at the start of day 1. This is in line with the literature that was mentioned in section 1.6 of Chapter 1 that once adults learn the new words, they often retain the information for a long period of time, even up to a year (e.g. Salasoo et al., 1985).

2.6.4 Chapter 5: Visual word learning in adults with dyslexia

Chapter 5 was intended to bring the learning paradigm that was developed in Experiment 3 of Chapter 2 to a more applied context in order to understand visual word learning in groups of dyslexic adults and normally-reading controls. The experimental design was exactly the same as those of Experiment 3 of Chapter 2. Result of Chapter 5 indicated that adults with dyslexia performed at a comparable level to typically-reading controls on a test of nonverbal ability (matrix reasoning) but had lower vocabulary scores, slower and less accurate reading and spelling of words, less efficient reading of nonwords, poorer phonological awareness, poorer performance on both verbal and nonverbal tests of span and working memory, and slower motor speed. These findings match other reports in the literature that dyslexics in higher education have cognitive problems that extend beyond reading and writing to wider aspects of linguistic, working memory and motor performance while having intact nonverbal reasoning (e.g. Hatcher et al., 2002; Warmington et al., 2013b). The largest difference between dyslexics and controls in Experiment 7 was on the TOWRE Phonemic Decoding Efficiency test (Torgesen et al., 1999), a test of nonword reading.

In the experimental task, the result of the control was very similar to those that were observed in Experiment 3 of Chapter 2. This section would focus on the result of the adults with dyslexia. Result of Experiment 7 showed that adults with dyslexia were capable of visual word learning. Yet, not only they were substantially slower at reading the nonwords throughout both sessions of the experiment compared to typical adults, they were also slower to build lexical entries in the orthographic and phonological lexicons. Whereas the difference in RTs between shorter and longer nonwords became nonsignificant in the typical readers around the middle of session 1, the dyslexics showed slower naming of longer nonwords throughout session 1, only losing the length effect part-way through session 2 (day 7).

Ability at reading and spelling real words ('literacy') predicted decoding speed across the two groups. When the effect of literacy was taken into account there was no additional effect of vocabulary, phonological awareness or working memory on decoding speed for these particular readers. The fact that Experiment 4 of Chapter 3 found a significant effect of phonological awareness in nonwords reading while Experiment 7 did not may due to the spelling task being included in Experiment 7 but not Experiment 4, and it may be the case that knowledge of the links between letters and sounds may be better captured by the kind of measures of word reading and spelling that went into the Literacy variable in Experiment 7.

Across the two groups, the ability to learn novel words (measured here as the change in RTs to longer nonwords between blocks 1 and 10 of day 1) was predicted by vocabulary and working memory. The finding that vocabulary predicts the ability to learn novel words was consistent with previous studies including Ricketts et al. (2009) and Storkel et al. (2006) and the result of Experiment 2 in Chapter 3. As regards the contribution of working memory, studies of children and young adults by Jarrold et al. (2009), Majerus and Boukebza (2013) and Martin and Ellis (2012) found a relationship between working memory and the ability to learn novel words, with working memory apparently related more closely to acquiring new word-forms rather than their meanings.

In terms of the DRC model of reading (Coltheart et al., 2001), poor reading of nonwords in the TOWRE-PDE test and in the experimental task reflects that adults with dyslexia have less efficient functioning of the nonlexical route compared to typical readers. Slower convergence between RTs to short and long nonwords (the length effect) in the dyslexics also suggest they required more exposure to the novel items before they can build representations in the orthographic input and phonological output lexicons.

This results in a slower transition from sub-lexical to predominantly lexical reading in the dyslexics. Lastly, the fact that nonword reading remains slower in the dyslexics than the controls even at the end of session two, combined with the fact that they read the real words in TOWRE–SWE slower than typical adults, indicates that the lexical route also functions less efficiently in adult dyslexics than in typical readers.

2.6.5 Chapter 6: Visual word learning in Chinese native speakers

Chapter 6 was intended to bring the learning paradigm that was developed in Experiment 3 of Chapter 2 to examine the process of visual English word learning in Chinese native speakers who speak Cantonese as their first language. The experimental design was exactly the same as those of Experiment 3 of Chapter 2. Experiment 9 mainly focuses on two questions in second-language acquisition. Firstly, do Chinese native speakers transfer their holistic approach in reading a logographic language (traditional Chinese characters) to reading an alphabetic language (English)? Secondly, given that Chinese character recognition does not implicitly require any phonological awareness (e.g. Siok & Fletcher, 2001) and English word reading in Hong Kong is taught using a ‘look and say’ method. This means that People in Hong Kong are not explicitly taught the instruction of phonological awareness. This may imply that Chinese native speakers may have a weaker grapheme-phoneme correspondence compared to British native speakers. The research question is whether this approach of learning English would guide Chinese readers to read English serially rather than lexically?

Result of Experiment 9 showed that the Cantonese native speakers performed at a comparable level to typically-reading British controls that Cantonese Chinese readers demonstrated a very high accuracy while reading the nonwords aloud. Cantonese Chinese speakers were clearly capable of visual word learning. Their naming latencies declined across blocks and their naming latencies to 4- and 7-letter nonwords eventually converged. Yet, learning occurred noticeably more slowly than in the British readers. Whereas the difference in RTs between shorter and longer nonwords became nonsignificant in the British readers around the middle of session 1, the Chinese readers showed slower naming of longer nonwords through session 1, only losing the length effect in the middle of session 2 (day 7). This experiment confirms that the problems with learning new English written words inefficiently in Cantonese Chinese speakers persist to early adulthood, even in highly literate adults (e.g. Cheung, 1995,1999; McBride-Chang et al., 2006).

Cantonese Chinese speakers were substantially slower at reading the nonwords throughout both sessions of the experiment. Importantly, the naming latencies for the Chinese speakers remained substantially longer than those of the British readers through to the end of session 2. This is consistent with the literature that Cantonese Chinese native speakers learn new written words slower than Mandarin and English native speakers (Holm & Dodd, 1996; Cheung et al; 2001). This can be explained by the phonological core view model in English word learning. The phonological core view model focuses on the specific linguistic skills that are important distinctively to L1 and L2 (Geva & Wang, 2001). Bringing this concept to the result of Experiment 9, this would mean an emphasis on the skills that are important in L2 (English) acquisition, which is phonological awareness in English nonword reading. The Cantonese Chinese participants learn English from a ‘look-and-say’ method (Bialystok et al., 2005) since they were 6 years old, they did not have the opportunity to acquire phonemic awareness before they learned English and the ‘look-and-say’ learning instruction does not facilitate phonemic awareness at that stage either. As a result, they would process the nonwords inefficiently. This explained why Cantonese native speakers showed longer RTs when they read the novel words for the 1st time in Block 1.

Comparing the result of Experiment 9 to the result of adults with dyslexia in Experiment 8, both groups showed a similar pattern of result. Both group were equally slow in reading nonwords and only managed to start reading the nonwords in a wholistic way for the first time in session 2 (day 7). Based on the linear mixed effect modelling result of Experiment 8, that fact that it took 15 presentations of the novel words before Chinese native speakers could build lexical entries in the orthographic and phonological lexicons may be due to their low level of vocabulary ability in English (as learning improvement was predicted by expressive vocabulary in Experiment 8).

7.2 Connectionist model and the Connectionist Dual Process (CDP++) model

This thesis has utilized the DRC model (Coltheart et al., 2001) to be the main framework to explain the reduction of length effect that was observed in Experiments 1 to 9. A major difference of the DRC model and the Connectionist model (Plaut et al., 1996) is that the DRC model does not contain a neural-net learning algorithm to a training set of stimuli, and it is often specified by the modeller on the basis of the empirical effects that the model is meant to explain (Snowling & Hulme, 2008). This section will include a brief description as to how the reduction of length effect that was observed in Experiment 1 to 9 can potentially be explain by the Connectionist model (Plaut et al., 1996) and Connectionist Dual Process (CDP++, Perry, Ziegler, & Zorizi, 2007; 2010) model. The main conclusion of this thesis is that the DRC model (Coltheart et al., 2001), Connectionist model (Plaut et al., 1996) and Connectionist Dual Process model (Perry et al., 2007; 2010) can explain the reduction of the length effect in Experiment 1 – 9 equally well.

2.6.6 Connectionist model

Connectionist models (Plaut et al., 1996) of word learning have been established as models of how the brain may learn and hold information. They contain a network of processing units (like neurons) that learn through experience with written words and feedback, via translating written words into sound and accessing their meanings. The triangle model is one type of the connectionist models, as shown in Figure 7.1. There are two distinct routes from orthography to phonology, one direct and the other through semantics. The indirect route (semantic route) from orthography to phonology is required when reading words aloud: it failed in reading nonwords. The direct route is required when reading nonwords aloud. Similarly, there are explicitly two routes from orthography to semantics, one direct and the other through phonology.

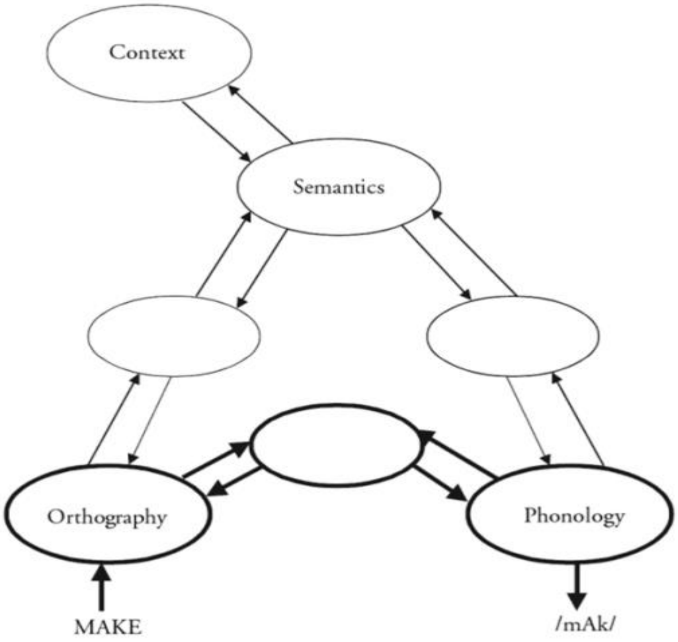

***Figure 7.1*. The general framework of the Connectionist model based on Seidenberg and McClelland (1989) (taken from Plaut et al., 1996).**

2.6.7 How words are processed and read

Within the connectionist model, information on words is stored in a distributed manner over different information units that co-operate with the orthographic, phonological and semantic components of the model. These critical components/layers are the larger ovals in Figure 7.1. The smaller ovals in the diagram represent hidden units. These 'hidden' units mediate the computations between codes. They increase the range and complexity of problems (e.g. the mapping between spelling and pronunciation) the model can solve and hence increase the computational power of the model (Rumelhart, Hinton, & Williams, 1986). Sometimes a layer of 'clean-up' units is connected to the output layer to reduce the noise and improve the settling process. Since all the major levels are linked together by the mediation of the hidden units, activation in any main layers produces activation in related layers.

Unlike the DRC model which has separate feature- and word-level representations, the triangle model does not. Entities such as spellings, pronunciations or word meanings are coded as patterns of activation over units encoding featural primitives. The precise pattern that is activated can change depending on the availability of contextual information. In contrast to the DRC model, there are no units representing individual words; instead the units represent sub-lexical features (graphemes, phonemes, sememes). These units operate as parts of the representation of many different words. For instance, the orthographic form of /ove/ is a part of the representations for love, prove, and cove. The input of the written strings is a pattern of activation across the orthographic units. The activation then expands through the network, ending in a pattern of activation over the phonological units, which represent the pronunciation of the word.

The second major difference between the triangle and DRC model is that the triangle models have a neural-net learning algorithm to a training set of stimuli whereas the DRC model relies on the programming of the modeller. When each item is presented to the triangle model, it is fed through the network and the output is produced. The output is compared with the correct ‘target’ value and the difference between the two is calculated for each output unit. The squared differences are summed over all the output units to give an overall measure of the ‘error’ that the network has generated. The aim of learning is to reduce the overall level of error and the back-propagation procedure can specify how the weights of the network (e.g. the strengths of the connections between the units) should be modified gradually in order to reduce the error.

Length effects have been much less central to the evaluation of connectionist models which have focused instead on overall differences in the efficiency of reading words and nonwords, the effects of word frequency and spelling-sound consistency on word reading, and the ability to simulate the disorders of reading seen in developmental and acquired dyslexia. Plaut et al. (1996, p. 85) reported effects of both orthographic and phonological length on the behaviour of their model, arguing that, "Even though the network settles to a representation of the phonemes of a word in parallel the time it takes to do so increases with the length of the word". Monaghan and Ellis (2010) showed that the degree of error associated with different words in the trained Harm and Seidenberg (1999) model was predicted by letter length as well as by word frequency, consistency and number of orthographic neighbours (i.e., the number of other words in the model's training set that differ from a target word by a single letter). This was true of the performance of a version of the model that was trained on all the words in its vocabulary together from the outset (as in Harm & Seidenberg, 1999) and was also true of the performance of a version of the model that reproduced normal reading development more closely by being trained first on words from Grade 1 reading material followed by the addition of words from Grades 2, 3 and so on in a cumulative, interleaved fashion (whereupon the model showed effects of age / order of acquisition alongside the effects of length and other factors). Parallel processing models can, therefore, show effects of length caused by the fact that more connections are involved in processing longer than shorter words, introducing more error or a longer settling time into the performance of the model. Neither Plaut et al. (1996) nor Monaghan and Ellis (2010) compared length effects in their models specifically for words and nonwords. Perry et al. (2007) reported, however, that the "triangle model" of Seidenberg and McClelland (1989) provided a poor fit to Weekes's (1997) results, failing to show any signs of differential effects of length in the reading of words and nonwords.

2.6.8 Connectionist Dual Process (CDP++) model

The models of Perry et al. (2007; 2010) combined distributed processing principles into their nonlexical and lexical routes, but maintained a distinction between those two very different ways of converting orthography to phonology. Figure 7.2 shows the schematic description of the CDP+ model. Both CDP+ and CDP++ have successfully replicated the effect of frequency and length and the interaction between length and lexicality. The following sections will briefly include the differences of the DRC and the CDP+ model in the sub-lexical and lexical route.

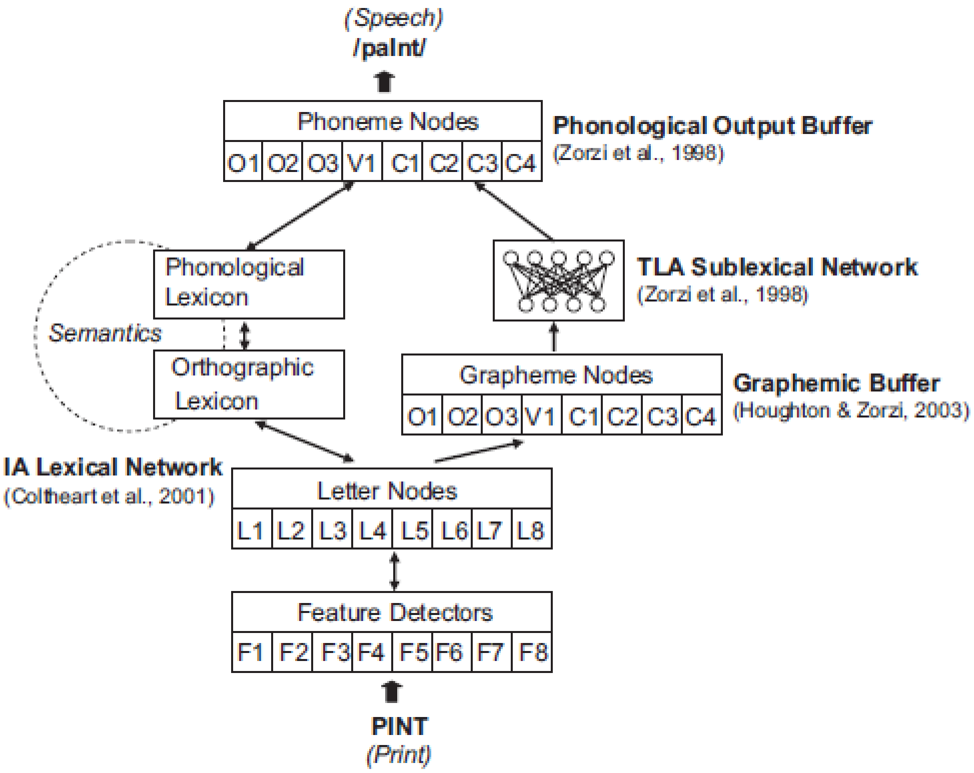

***Figure 7.2.* Schematic description of the Connectionist Dual Process model (CDP+) (taken from Perry et al., 2007)**

7.2.1.1 The sub-lexical route

An orthographic buffer was implemented in the sub-lexical route of the CDP+ model. Single input notes do not represent individual letters only, but also complex graphemes such as *ck*, *th*, etc. When letters combine to form one of these graphemes, the grapheme is activated instead of letters. The input representations then align graphemes into onset slots, vowel slots and coda slots. The phonological output of the network also includes the onset, vowel, and coda slots. Thus, when training patterns are presented to the network, the output (phonological response) is broken down into onset-vowel-coda.

7.2.1.2 The lexical route

The lexical route of the CDP+ model still includes the letter feature level, the letter level, an orthographic lexicon, a phonological lexicon and the phonological output buffer. In fact, the lexical route of the CDP+ model is identical to that of DRC all the way up to and including the phonological lexicon, excluding the null characters (in word coding, when a word is less than 8 letters, the null letter is put on to the end of all words to make it up to 8 positions). The phonological output buffer was also changed so instead of the phonemes being aligned as a contiguous string, the phonemes were aligned so that they follow the onset-vowel-coda distinction. Furthermore, the frequencies of the words in the phonological lexicon were changed so that they were phonological rather than orthographic frequencies (unlike the implementation of DRC model).

7.3 Alternative explanation (the power law)

A power law is often used to describe the effect of practice on learning and memory (Logan, 1992; Newell & Rosenbloom, 1981) and indicates an improvement in performance that reduces gradually over time. The power law can help to describe the pattern that is observed for the naming task in the thesis. For example, the Day 1 naming reaction time data (RTs) from Experiment 3 can be described by the negatively accelerated function of power law (Logan, 1990), that is

$$\mathrm{RT} = \mathrm{a} + \mathrm{bN}^{-\mathrm{c}}$$

RT is the reaction time, *N* is the number of practice trials. *A* is asymptote, reflecting the irreducible limit on performance. *b* is the difference between initial and asymptotic performance, reflecting the amount to be learned. *C* is the exponent, reflecting the rate of learning. Based on the aforementioned principles, the exponential power function for Day 1 4-letter item of Experiment 3 is, $RT = 498 + 54\mathrm{N}^{-0.014}$, and $RT = 510 + 123\mathrm{N}^{-0.025}$, for Day 1 7-letter items. By comparing the two formulas, the 7-letter nonwords showed 1) longer RTs for asymptote (the difference of *A* between the two formulas was 12); 2) a bigger difference between initial and asymptotic performance (the difference of *b* between the two formulas was 69) and 3) the rate of learning was larger (the difference of *c* was 0.011).

2.6.9 Are the characteristics of word learning similar to those in general skill acquisition?

Power functions have been utilized to describe the results of a wide range of tasks, including free recall of word list (Wixted & Ebbesen, 1991), learning lists of nonsense syllables (Wixted & Carpenter, 2007), and arithmetic skills acquisition (Delaney, Reder, Staszewski, & Ritter, 1998). There is a possibility that certain mechanisms that support word learning reflect the properties the system that supports general skills acquisition. If so, the convergence we see between longer and shorter sequences of letters in word learning might also be seen in other tasks in which participants learn sequences of items (e.g., faces or musical notes). Such sequences do not naturally elicit a single response the way that novel words elicit their names. In order to compare learning and unitisation for words with sequences of faces (for example) it would be necessary to change the task. One plausible way to explore this question would be to utilize a task in which participants must respond whenever they see a target letter or face within a sequence, with the same sequences being presented repeatedly across blocks. The 'word superiority effect' demonstrates that a single target letter is identified more efficiently in the context of a familiar word or well-structured nonword than within a randomly structured letter string or a letter in insolation (Reicher, 1969). I have not been able to identify any studies of letter-in-word tasks where the stimuli have been repeated to discover the effects of learning on performance. Future experiments could compare letters with faces and explore whether length effects in responses to target faces in short and long sequences will reduce when some of the sequences are repeated. If a similar effect was found for the target faces, this would suggest that the reduction of length effect in the naming task is similar to those in general skills acquisition.

Another possible way to understand this question further would to study note-reading in music performance. In music, notes are the functional analogue of letters. Note-reading comprises the translation from the visual domain to a representation which provides the information for a program specifying the patterning and timing. Sloboda, Clarke, Parncutt, and Raekallio (1998) found that pianists who used the same fingering strategies to play repeated sequence result in a higher speed and accuracy measure. This demonstrates that the reduction of RTs in the naming task maybe related to those that are observed in other skills acquisition.

The fact that certain characteristics of word learning may be similar to those in general skills acquisition does not undermine the purpose of this thesis. It would simply show that the formation of lexical entries is achieved by processes similar to those responsible for other forms of learning. In some ways, it would be surprising if they were not. This thesis aims to investigate the process of lexicalization by which new items are considered 'lexical' come into being. A broad definition of lexicalization can be found in Brinton (2002)--- the ordinary process of word formation in the lexicon. By designing nine experiments, this thesis has explored the process of orthographic and phonological processing in English word learning. Thus, the purpose of this thesis is valid even if there are certain characteristics of word learning that are similar to those in general skills acquisition.

2.6.10 Can the power law explain the reduction of length effect?

Originally the power law is not implemented to investigate the effect of length. Zoccolotti, De Luca, Di Filippo, Judica, and Martelli (2009) tried to understand whether the specific effects of length, word frequency and lexicality still significantly affect naming RTs even when the global factors in reading spend was examined (e.g reduced naming RTs through practice). They asked 503 first-to-eighth graders to read aloud Italian nonwords. By using the power law function, Zoccolotti et al. (2009) found that the global processing factor accounted for a large portion of the variance. Yet, specific influences of length, frequency and lexicality were detected over and above the global processing factor. Based on the aforementioned studies, the power law seems to be a helpful tool to describe the result in this thesis. Yet, it does not provide any explanation as to *why* there is the reduction of length is significant even when the general effect (e.g. practice effect) is taken into account.

7.4 Implications of the findings

Experiments 1 – 3 have developed a paradigm that has considerably potential as a tool for investigating visual word learning. By using the reduction of naming RTs and length effect, this gives an opportunity to understand when does unitization happens in the process of visual word learning. All 9 experiments in this thesis show that typical British adults learn new items rapidly. It only takes them around six exposures to build lexical entries in the orthographic and phonological lexicons. Furthermore, once these entries are built, the representations are very robust that the learned materials can be retained even up to a week (Experiment 3) and a month (Experiment 7) without further revision.

The results of Experiment 4 – 6 were in accordance to the ‘lexical quality hypothesis’ (Perfetti, 1992), learning was better in the read aloud training conditions which participants had the benefit to receive training in both the orthography and the phonology of the stimuli. Yet, simply by training the phonology of the novel words is sufficient to build lexical entries in the orthographic and phonological lexicons. Experiment 4 highlights the importance of literacy (word and nonword reading composite) and phonological awareness in reading nonwords. Vocabulary was crucial to the improvement of both orthographic and phonological learning while RDN was only crucial for orthographic learning. Experiment 6 shows that trained items were still learned better compared to untrained items even when distractors were included in the hear-and-repeat training in order to attenuate the activation of orthographic codes.

Experiment 7 showed that by interleaving high-, low-frequency and nonwords in 20 blocks of reading aloud task, the word frequency effect was modulated by recent experience. Though there was some slowing down when the participants named the nonwords when they were back on day 28, it only took participants one block before they could read the words and nonwords as quickly as the end of day 1. This again shows that British typical adults have good retention of learned materials.

Experiment 8 shows that adult dyslexics in higher education continue to experience difficulty reading novel words and nonwords. They are slower to read nonwords aloud than typical readers, requiring more time per letter to pronounce unfamiliar sequences of letters. Though they are capable of visual learning, they require more exposures than typical readers before they establish effective lexical representations. They remain slower than typical readers even at reading familiar words aloud. Across both dyslexic and typical readers, decoding speed for nonwords was predicted by skill at reading and spelling real words ('literacy') while individual differences in word learning were predicted by vocabulary size and working memory.

Result of Experiment 8 showed that the problems that adult dyslexics experience extend beyond reading and spelling to word learning, vocabulary, phonological awareness, working memory and even basic motor speed. Taken together, those problems will conspire to make it very challenging for adult dyslexics to function successfully within higher education.

Experiment 9 showed that the relatively weak nonword reading ability in Cantonese native speakers persists to adulthood. Along with the results in the literature (e.g. Holm & Dodd, 1996), the result of Chapter 9 also showed that while phonological awareness is not a prerequisite for the achievement of high levels of literacy in English, training in phonological awareness allows the use of phonics for learning new words efficiently. Thirdly, while the rich oral language environment may help to develop a very fundamental understanding towards phoneme segmentation, a more sophisticated level of phonological awareness requires explicit instruction for development. Finally, from an educational point of view, pre-sessional English courses that are tailored made for Chinese students to prepare them for high-education should include training in both phonological awareness and vocabulary. Despite being highly literate, the fact that Cantonese native speakers do not receive any explicit phonological awareness training and this may imply that they may demonstrate similar learning difficulties as with British students with dyslexia (e.g. spelling problems), which may poses a problem when they have to hand in written assignments and sit for exams in higher education.

7.5 Future directions

There are, as always, issues remaining to be resolved. Firstly, certain variables can be controlled better. One of these is the extent to which the differences in neighbourhood density between shorter and longer nonwords contribute to the effects that were found in this thesis and in the literature. Another is the relative contribution of orthographic and phonological length (including number of syllables) to the length effects observed in this thesis. Forster and Chambers (1973) found an effect of length on naming latencies measured across both words and nonwords that appeared to be linked to the number of letters in the stimuli rather than the number of syllables. In contrast, a large-scale study by Yap and Balota (2009) found independent effects of both number of letters and number of syllables on word naming latencies, suggesting that both orthographic and phonological length may play a role in determining word naming latencies, with those effects probably being driven by lower frequency words.

New, Ferrand, Pallier, and Brysbaert (2006) reported that the effect of length on word naming latency is nonlinear, with mean naming latencies reducing between 3 and 5 letters then increasing from 7 or 8 letters onwards. Mean naming latencies for nonwords in Maloney et al. (2009), Weekes (1997) and all 9 experiments of the present thesis show naming latencies increase with letter length from 3 letters upwards. A rather different pattern was reported by New et al. (2006) for word naming, with naming latencies declining as length increases from 3 to 5 letters then increasing beyond that point. This prompts the question of whether lexicalisation changes the shape of the function relating letter length to naming latencies for (lower frequency) words compared with nonwords. The high- and low-frequency words in this thesis were also not matched on AoA. Future research can try to separate these factors, establishing for example if length effects are greater for late than early acquired words.

Secondly, the blocking effect observed in Experiment 2 still required further investigation. Lupker et al. (1997) and Rastle et al. (2003) observed that high-frequency words are named more slowly when mixed with nonwords than when presented in unmixed ("pure") blocks of trials while nonwords showed the opposite pattern, being named more slowly when presented on their own in pure blocks than when interleaved with high-frequency words in mixed blocks of trials. Lupker et al. (1997) and Rastle et al. (2003) argued that participants set a criterion for the speed of responding to stimuli in a block based on the blend of easy or difficult items within the block. When the items are all easy (e.g., pure blocks of high-frequency words) the criterion will be relatively short and RTs consequently faster. When the items are a mixture of easy and difficult, a criterion will be set that is somewhere between resulting in a homogenization of RTs to easier and more difficult items. That was not the pattern seen in the present Experiment 2 where untrained (difficult) nonwords gained in block 10 from being mixed with trained (easy) nonwords but RTs to the trained (easy) nonwords were barely affected (if at all) by being mixed with new, difficult nonwords. Thus, further studies can try to understand the discrepancy between the result of Experiment 2 and those in the literature.

Thirdly, this thesis did not pay enough attention to understand how the learned materials can be retained for a month without further revision. This is related to the "complementary learning systems" approach to learning presented by McClelland et al. (1995) and applied to word learning by Davis and Gaskell (2009). The complementary learning systems approach proposes that when new connections must be created between representations in different parts of the brain (e.g., the orthographic and phonological representations of novel words), the hippocampus and associated cortex is initially involved in forging those connections. Over time, and as a result of consolidation processes that may be facilitated by sleep (e.g., Tamminen, Payne, Stickgold, Wamsley, & Gaskell, 2010), those connections are established at a purely cortical level, freeing the hippocampus for new learning. O'Reilly et al. (2011) argued that consolidation and transfer of information to the cortex helps protect against interference. A hallmark of the transfer from hippocampal to cortical connections is the emergence of competition effects between newly-learned words and established vocabulary (e.g., Henderson, Weighall, Brown, & Gaskell, 2013). If so, then under the conditions of the present experiments, future studies may find competition between novel written words and established words in the lexicon of the sort reported by Bowers, Davis, and Hanley (2005) should be observed after a period of consolidation (e.g., session 2 of the present Experiment 3), but not within the initial learning session.

Fourthly, the possible additional impact of associating meanings with the novel words, as happens in natural language learning (cf. McKague et al., 2008; McKague et al., 2001) could also be included. McKague et al. (2008) found reading aloud performance became faster and more accurate for novel words in the semantic condition, but only for novel words with inconsistent pronunciations. This semantic advantage for inconsistent novel words was again observed when a subset of participants were retested 6 – 12 months later. Thus, future studies can investigate whether the speed of lexicalization can be improved by incorporating meanings in the novel words.

Fifthly, future studies can explore whether there is a benefit from phonological training to reading nonwords that share the same consonants but have different vowels. The result of Experiments 4 – 6 are in line with Johnston et al. (2004) and McKague et al. (2008) that at least a briefly formed orthographic representation of the novel words were encoded in the hear-and-repeat training prior to the first visual encounter of the novel words. Based on McKague et al. (2008), the orthographic representation that was generated by the phonological training would mainly be framed by the consonants of the novel words. If McKague et al.'s (2008) suggestion is right, then one would expect that haing phonological training on the nonword *blispod* would reduce the naming RTs of *blespud* as the two nonwords share the same consonants but not the vowels. Moreover, Experiment 4 to 6 also did not include the equivalent condition that solely involves orthographic learning that does not involve the activation by the phonological codes. Future studies can address this issue by adopting an artificial orthography paradigm that utilized novel characters (Taylor et al., 2011).

Sixthly, regarding the cognitive assessments that were used in Experiment 4 and Experiment 8, only one task was used in each element that taps into the cognitive skills of the participants. Future studies should incorporate a few more tests in each predictor. For example, different tests that tap into phonological awareness in different levels should be included, including the simple phoneme deletion task and the more difficult ones (e.g. spoonerism). This may show that people with weaker phonological ability, e.g. adults with dyslexia and adults who speak English as a second language (ESL), may be able to complete the basic phoneme deletion task but not the spoonerism task. Furthermore, based on the result of Experiment 8, vocabulary was the important predictor of English word learning. Chinese students presumably have a large vocabulary size in Chinese, but a smaller English vocabulary than English native speakers. This suggests that vocabulary size in English is what predicts English word learning. This can be tested in a future experiment by measuring L1 and L2 vocabularies separately. Furthermore, the cognitive profile of the participants was not recorded in Experiment 9. Thus, is it not certain what cognitive skills do Hong Kong readers utilize to help them to learn English (e.g. working memory). Future studies should include these tasks in order to pinpoint what specific cognitive skills are important in second language acquisition.

Finally, one application would be to study how the effect of length differs in various language groups; for example, Spanish (Ferrand, 2000) and Chinese (Ho & Bryant, 1997). Avdyli, Kwok, Bermudez, Cuetos and Ellis (in preparation) found that when Spanish speakers read aloud Spanish nonwords, the effect of length persisted even when participants read the novel words aloud for 20 times. This may be related to the fact that Spanish is much more phonologically transparent than English. Therefore, the simpler grapheme-phoneme conversion rules of Spanish mean that the nonlexical conversion would contribute more to reading aloud in Spanish than in English, resulting in length effects even participants has seen the novel words for many times. If word learning differs across different orthographies, it would indicate that any domain-general aspects of word learning are modulated by language-specific characteristics. The descriptive value of the power law still needs to be accompanied by functional explanations that take account of different language characteristics.

7.6 Conclusions

The research reported in this thesis adds to a growing body of research suggesting that the process of visual word learning in adults occurs rapidly. Furthermore, once these lexical entries are formed in the mental lexicon, the word representations are resilient enough that it can be retained in memory for up to a month without further revision. In accordance to the ‘lexical quality hypothesis’ (Perfetti, 1992), learning is always better when both the input of orthography and phonology is available. Yet, just by training the phonology of a novel word is sufficient to build lexical entries in the lexicon.

This thesis has also confirmed that the problems that have been documented in dyslexic children persist into early adulthood, even in high-functioning dyslexics. Not only adults with dyslexia were a lot slower in reading novel words, they were also slower to reach parallel reading compared to typical control. Given that the behavioural result of adult dyslexics and second language speakers were so similar, it is hoped that the findings of the studies reported here will inform interventions for dyslexics and pre-sessional English courses for Chinese native speakers.

1 Appendix 1

1.1 Analysis of the results in Experiments 1, 2, and 3 of Chapter 2.

1.1.1 Experiment 1

Effects	*By-participants analysis*	*By-items analysis*
NAMING LATENCY (correct, trimmed RTs)		
Overall analysis		
Length	F_1 (1, 24) = 51.95, MSE = 110589, p <.001, η_p^2 = .684	F_2 (1, 22) = 11.15, MSE = 52068, p <.005, η_p^2 = .336
Blocks	F_1 (3.50, 83.97) = 38.45, MSE = 263970, p < .001, η_p^2 = .616	F_2 (9, 198) = 93, MSE = 49174, p < .001, η_p^2 = .809
Length x Blocks	F_1 (4.72, 113.16) = 8.63, MSE = 14197, p < .001, η_p^2 = .264	F_2 (9, 198) = 6.81, MSE = 3602, p < .001, η_p^2 = .236
Post-hoc comparison of RTs to 4- and 7-letter items in blocks 1-10 using Bonferroni-corrected t-tests ($\propto$ = .005)		
Block 1	t_1(24) = 5.20, p < .001	t_2(24) = 20.15, p < .001
Block 2	t_1 (24) = 6.77, p < .001	t_2 (24) = 27.50, p < .001
Block 3	t_1 (24) = 4.69, p < .001	t_2 (24) = 9.74, p = .005
Block 4	t_1 (24) = 1.88, p = .070	t_2 (24) = 1.33, p = .261
Block 5	t_1 (24) = 3.51, p < .005	t_2 (24) = 4.92, p = .037
Block 6	t_1(24) = 3.89, p < .001	t_2(24) = 3.59, p = .071
Block 7	t_1 (24) = 1.26, p = .222	t_2 (24) = 0.57, p = .460
Block 8	t_1 (24) = 1.45, p = .161	t_2 (24) = 1.15, p = .295
Block 9	t_1 (24) = 1.88, p = .072	t_2 (24) = 1.38, p = .253
Block 10	t_1 (24) = 1.70, p = .101	t_2 (24) = 1.49, p = .236

1.1.2 Experiment 2

Effects	*By-participants analysis*	*By-items analysis*
Analysis of trained items only		
NAMING LATENCY (correct, trimmed RTs)		
Overall analysis		
Length	$F_1(1, 23) = 24.78$, $MSE = 115165$, $p < .001$, $\eta_p^2 = .519$	$F_2(1, 22) = 19.22$, $MSE = 59158$, $p < .001$, $\eta_p^2 = .466$
Blocks	$F_1(2.41, 55.37) = 28.52$, $MSE = 309800$, $p < .001$, $\eta_p^2 = .554$	$F_2(9, 198) = 113.23$, $MSE = 41932$, $p < .001$, $\eta_p^2 = .837$
Length x Blocks	$F_1(3.76, 86.46) = 12.46$, $MSE = 24623$, $p < .001$, $\eta_p^2 = .351$	$F_2(9, 198) = 13.86$, $MSE = 5133$, $p < .001$, $\eta_p^2 = .387$
Post-hoc comparison of RTs to 4- and 7-letter items in blocks 1-10 using Bonferroni-corrected t-tests		
(∝ = .005)		
Block 1	$t_1(23) = 6.01$, $p < .001$	$t_2(23) = 49.51$, $p < .001$
Block 2	$t_1(23) = 4.09$, $p < .001$	$t_2(23) = 17.27$, $p < .001$
Block 3	$t_1(23) = 2.57$, $p = .017$	$t_2(23) = 9.37$, $p < .01$
Block 4	$t_1(23) = 3.20$, $p < .005$	$t_2(23) = 9.04$, $p < .01$
Block 5	$t_1(23) = 3.22$, $p < .005$	$t_2(23) = 6.16$, $p = .021$
Block 6	$t_1(23) = 2.02$, $p = .055$	$t_2(23) = 3.02$, $p = .096$
Block 7	$t_1(23) = 1.73$, $p = .098$	$t_2(23) = 1.40$, $p = .250$
Block 8	$t_1(23) = .53$, $p = .605$	$t_2(23) = .16$, $p = .692$
Block 9	$t_1(23) = 3.02$, $p = .006$	$t_2(23) = 4.68$, $p = .042$
Block 10	$t_1(23) = 2.65$ $p = .014$	$t_2(23) = 9.63$, $p < .005$

Comparison of Trained and Untrained NAMING LATENCY (correct, trimmed RTs)		
Overall analysis		
Training	$F_1(1, 23) = 16.64$, $MSE = 24911$, $p < .001$, $\eta_p^2 = .420$	$F_2(1, 22) = 35.11$, $MSE = 12490$, $p < .001$, $\eta_p^2 = .615$
Length	$F_1(1, 23) = 41.84$, $MSE = 313714$, $p < .001$, $\eta_p^2 = .645$	$F_2(1, 22) = 61.42$, $MSE = 159985$, $p < .001$, $\eta_p^2 = .736$
Blocks	$F_1(1, 23) = 37.51$, $MSE = 325464$, $p < .001$, $\eta_p^2 = .620$	$F_2(1, 22) = 137.26$, $MSE = 162279$, $p < .001$, $\eta_p^2 = .862$
Train x Length	$F_1(1, 23) = 5.55$, $MSE = 11148$, $p < .05$, $\eta_p^2 = .194$	$F_2(1, 22) = 16.96$, $MSE = 6033$, $p < .001$, $\eta_p^2 = .435$
Train x Blocks	$F_1(1, 23) = 18.57$, $MSE = 57167$, $p < .001$, $\eta_p^2 = .447$	$F_2(1, 22) = 35.24$, $MSE = 31501$, $p < .001$, $\eta_p^2 = .616$
Length x Blocks	$F_1(1, 23) = 25.04$, $MSE = 39704$, $p < .001$, $\eta_p^2 = .521$	$F_2(1, 22) = 15.84$, $MSE = 18732$, $p < .001$, $\eta_p^2 = .419$
Train x Length x Blocks	$F_1(1, 23) = 7.47$, $MSE = 7563$, $p < .05$, $\eta_p^2 = .245$	$F_2(1, 22) = 4.44$, $MSE = 3966$, $p < .05$, $\eta_p^2 = 168$
4-letter		
Training	$F_1(1, 23) = 1.68$, $MSE = 1365$, $p = .208$, $\eta_p^2 = .068$	$F_2(1, 11) = 2.44$, $MSE = 581$, $p = .146$, $\eta_p^2 = .182$
Blocks	$F_1(1, 23) = 23.84$, $MSE = 68908$, $p < .001$, $\eta_p^2 = .509$	$F_2(1, 11) = 45.93$, $MSE = 35371$, $p < .001$, $\eta_p^2 = .807$
Train x Blocks	$F_1(1, 23) = 11.59$, $MSE = 11573$, $p < .005$, $\eta_p^2 = .335$	$F_2(1, 11) = 12.42$, $MSE = 6557$, $p < .005$, $\eta_p^2 = .530$

Post-hoc comparison of RTs to trained and untrained 4-letter nonwords in Block 1 and 10 using t-tests (∝ = .05)		
Block 1	$t_1(23) = 1.64$, $p = .115$	$t_2(11) = 1.89$, $p = .085$
Block 10	$t_1(23) = 3.44$, $p < .005$	$t_2 (11) = 4.19$, $p < .005$
7-letter		
Training	$F_1(1, 23) = 12.88$, $MSE = 34694$, $p < .005$, $\eta_p^2 = .359$	$F_2(1, 11) = 37.88$, $MSE = 17941$, $p < .001$, $\eta_p^2 = .775$
Blocks	$F_1(1, 23) = 40.19$, $MSE = 296259$, $p < .001$, $\eta_p^2 = .636$	$F_2(1, 11) = 91.34$, $MSE = 145640$, $p < .001$, $\eta_p^2 = .893$
Train x Blocks	$F_1(1, 23) = 17.19$, $MSE = 53157$, $p < .001$, $\eta_p^2 = 428$	$F_2(1, 11) = 22.94$, $MSE = 28910$, $p < .001$, $\eta_p^2 = .676$
Post-hoc comparison of RTs to trained and untrained 7-letter nonwords in Block 1 and 10 using t-tests (∝ = .05)		
Block 1	$t_1(23) = 0.63$, $p = .534$	$t_2(11) = 0.91$, $p = .383$
Block 10	$t_1(23) = 5.11$, $p < .001$	$t_2(11) = 7.00$, $p < .001$
Analysis of RTs to trained items in Blocks 1 and 10		
Length	$F_1(1, 23) = 32.54$, $MSE = 103294$, $p < .001$, $\eta_p^2 = .586$	$F_2(1, 22) = 51.12$, $MSE = 51943$, $p < .001$, $\eta_p^2 = .699$
Blocks	$F_1(1, 23) = 42.29$, $MSE = 327718$, $p < .001$, $\eta_p^2 = .648$	$F_2 (1, 22) = 224.58$, $MSE = 168389$, $p < .001$, $\eta_p^2 = .911$
Length x Blocks	$F_1(1, 23) = 25.12$, $MSE = 40961$, $p < .001$, $\eta_p^2 = .522$	$F_2 (1, 22) = 26.63$, $MSE = 19968$, $p < .001$, $\eta_p^2 = .548$

Analysis of RTs to untrained items in Blocks 1 and 10		
Length	$F_1(1, 23) = 34.98$, $MSE = 221568$, $p < .001$, $\eta_p^2 = .603$	$F_2(1, 22) = 58.67$, $MSE = 114075$, $p < .001$, $\eta_p^2 = .727$
Blocks	$F_1(1, 23) = 13.71$, $MSE = 54913$, $p < .001$, $\eta_p^2 = .373$	$F_2(1, 22) = 19.14$, $MSE = 25392$, $p < .001$, $\eta_p^2 = .465$
Length x Blocks	$F_1(1, 23) = 6.52$, $MSE = 6305$, $p < .05$, $\eta_p^2 = .221$	$F_2(1, 22) = 2.06$, $MSE = 2730$, $p = .165$, $\eta_p^2 = .086$

1.1.3 Experiment 3

Effects	*By-participants analysis*	*By-items analysis*
NAMING LATENCY (correct, trimmed RTs)		
Overall analysis		
Day	$F_1(1, 39) = 9.73$, $MSE = 295066$, $p < .005$, $\eta_p^2 = .200$	$F_2(1, 22) = 164.22$, $MSE = 89637$, $p < .001$, $\eta_p^2 = .882$
Length	$F_1(1, 39) = 41.28$, $MSE = 133023$, $p < .001$, $\eta_p^2 = .514$	$F_2(1, 22) = 12.09$, $MSE = 39470$, $p < .005$, $\eta_p^2 = .355$
Blocks	$F_1(2.91, 113.37) = 60.79$, $MSE = 445204$, $p < .001$, $\eta_p^2 = .609$	$F_2(9, 198) = 81.83$, $MSE = 43961$, $p < .001$, $\eta_p^2 = .788$
Day x Length	$F_1(1, 39) = 34.41$, $MSE = 44097$, $p < .001$, $\eta_p^2 = .469$	$F_2(1, 22) = 24.08$, $MSE = 13142$, $p < .001$, $\eta_p^2 = .523$
Length x Blocks	$F_1(5.18, 201.99) = 24.46$, $MSE = 24550$, $p < .001$, $\eta_p^2 = .385$	$F_2(9, 198) = 8.09$, $MSE = 4344$, $p < .001$, $\eta_p^2 = .269$
Day x Blocks	$F_1(3.96, 154.24) = 16.68$, $MSE = 65356$, $p < .001$, $\eta_p^2 = .300$	$F_2(9, 198) = 42.08$, $MSE = 8726$, $p < .001$, $\eta_p^2 = .657$
Day x Length x Blocks	$F_1(5.10, 198.86) = 7.62$, $MSE = 7571$, $p < .001$, $\eta_p^2 = .163$	$F_2(9, 198) = 6.44$, $MSE = 1336$, $p < .001$, $\eta_p^2 = .227$

Day 1		
Length	$F_1(1, 39) = 52.86$, $MSE = 165150$, $p < .001$, $\eta_p^2 = .575$	$F_2(1, 22) = 18.62$, $MSE = 49082$, $p < .001$, $\eta_p^2 = .458$
Blocks	$F_1(2.84, 110.78) = 65.14$, $MSE = 473787$, $p < .001$, $\eta_p^2 = .626$	$F_2(9, 198) = 104.61$, $MSE = 45623$, $p < .001$, $\eta_p^2 = .826$
Length x Blocks	$F_1(3.60, 140.56) = 26.70$, $MSE = 41536$, $p < .001$, $\eta_p^2 = .406$	$F_2(9, 198) = 11.76$, $MSE = 5129$, $p < .001$, $\eta_p^2 = .348$
Simple main effects analyses of overall data		
Blocks at 4 letters	$F_1(9, 31) = 9.50$, $MSE = 0.73$, $p < .001$, $\eta_p^2 = .734$	$F_2(9, 14) = 7.60$, $MSE = 0.83$, $p < .001$, $\eta_p^2 = .830$
Blocks at 7 letters	$F_1(9, 31) = 19.59$, $MSE = 0.85$, $p < .001$, $\eta_p^2 = .850$	$F_2(9, 14) = 29.63$, $MSE = 0.95$, $p < .001$, $\eta_p^2 = .950$
Post-hoc comparison of RTs to 4- and 7-letter items in blocks 1-10 using Bonferroni-corrected t-tests (∝ = .005)		
Block 1	$t_1(39) = 8.42$, $p < .001$	$t_2(23) = 42.31$, $p < .001$
Block 2	$t_1(39) = 6.27$, $p < .001$	$t_2(23) = 13.91$, $p = .001$
Block 3	$t_1(39) = 4.47$, $p < .001$	$t_2(23) = 7.26$, $p = .013$
Block 4	$t_1(39) = 3.26$, $p < .005$	$t_2(23) = 2.47$, $p = .131$
Block 5	$t_1(39) = 5.26$, $p < .001$	$t_2(23) = 10.42$, $p = .004$
Block 6	$t_1(39) = 2.60$, $p = .014$	$t_2(23) = 1.87$, $p = .186$
Block 7	$t_1(39) = 4.12$, $p < .001$	$t_2(23) = 4.05$, $p = .057$
Block 8	$t_1(39) = 1.98$, $p = .055$	$t_2(23) = 0.82$, $p = .376$
Block 9	$t_1(39) = 2.14$, $p = .039$	$t_2(23) = 1.96$, $p = .175$
Block 10	$t_1(39) = 1.91$, $p = .063$	$t_2(23) = 2.17$, $p = .155$

Day 7		
Length	$F_1(1, 39) = 8.68$, *MSE* = 11971, $p < .005$, $\eta_p^2 = .182$	$F_2(1, 22) = 3.01$, *MSE* = 3531, $p = .097$, $\eta_p^2 = .120$
Blocks	$F_1(3.64, 141.90) = 12.83$, *MSE* = 56872, $p < .001$, $\eta_p^2 = .248$	$F_2(9, 198) = 22.91$, *MSE* = 7064, $p < .001$, $\eta_p^2 = .510$
Length x Blocks	$F_1(6.42, 250.42) = 3.45$, *MSE* = 2500, $p < .005$, $\eta_p^2 = .081$	$F_2(9, 198) = 1.79$, *MSE* = 551, $p = .073$, $\eta_p^2 = .075$
Simple main effects analyses of overall data		
Blocks at 4 letters	$F_1(9, 31) = 3.81$, *MSE* = 0.53, $p < .005$, $\eta_p^2 = .525$	$F_2(9, 14) = 11.21$, *MSE* = 0.88, $p < .001$, $\eta_p^2 = .878$
Blocks at 7 letters	$F_1(9, 31) = 6.25$, *MSE* = 0.65, $p < .001$, $\eta_p^2 = .645$	$F_2(9, 14) = 32.71$, *MSE* = 0.96, $p < .001$, $\eta_p^2 = .955$
Post-hoc comparison of RTs to 4- and 7-letter items in blocks 1-10 using Bonferroni-corrected t-tests (∝ = .005)		
Block 1	$t_1(39) = 4.35$, $p < .001$	$t_2(23) = 13.31$, $p < .001$
Block 2	$t_1(39) = 1.70$, $p = .098$	$t_2(23) = 0.95$, $p = .341$
Block 3	$t_1(39) = 1.89$, $p = .066$	$t_2(23) = 1.62$, $p = .217$
Block 4	$t_1(39) = 0.18$, $p = .855$	$t_2(23) = 0.01$, $p = .962$
Block 5	$t_1(39) = 1.23$, $p = .227$	$t_2(23) = 0.55$, $p = .467$
Block 6	$t_1(39) = 0.15$, $p = .882$	$t_2(23) = 0.07$, $p = .790$
Block 7	$t_1(39) = 0.20$, $p = .844$	$t_2(23) = 0.08$, $p = .930$
Block 8	$t_1(39) = 0.02$, $p = .983$	$t_2(23) = 0.01$, $p = .977$
Block 9	$t_1(39) = 2.00$, $p = .053$	$t_2(23) = 1.73$, $p = .202$
Block 10	$t_1(39) = 1.43$, $p = .160$	$t_2(23) = 0.61$, $p = .442$

Comparison between Day 1 Block 10 and Day 7 Block 1 (ANOVAs)		
Length	F_1(1, 39) = 15.52, *MSE* = 19139, $p < .001$, η_p^2 = .285	F_2(1, 22) = 11.82, *MSE* = 5466, $p < .005$, η_p^2 = .350
Blocks	F_1(1, 39) = 19.75, *MSE* = 99939, $p < .001$, η_p^2 = .336	F_2(1, 22) = 93.15, *MSE* = 29601, $p < .001$, η_p^2 = 809
Length x Blocks	F_1(1, 39) = 6.32, *MSE* = 3925, $p < .05$, η_p^2 = .139	F_2(1, 22) = 4.22, *MSE* = 1340, $p = .052$, η_p^2 = .161
Comparison between 4- and 7-letter within Day 1Block 10 and Day 7 Block 10		
Block 1	t_1(39) = 1.91, $p = .063$	t_1(23) = 2.17, $p < .155$
Day 1 Block 10	t_1(39) = 4.35, $p < .001$	t_1(23) = 13.31, $p < .001$
Day 7 Block 1		
Comparison between Day 1 Block 10 and Day 7 Block 1		
4-letter	t_1(39) = 3.34, $p < .005$	t_2(23) = 4.75, $p < .001$
7-letter	t_1(39) = 5.06, $p < .001$	t_2(23) = 9.74, $p < .001$

2 Appendix 2

2.1 Analysis of the results in Experiments 4, 5, and 6 of Chapter 3.

2.1.1 Experiment 4

Effects	*By-participants analysis*	*By-items analysis*
NAMING LATENCY (correct, trimmed RTs)		
Trained items(Block 1 & 10)		
Blocks	$F_1(1, 39) = 14.10$, $MSE = 176720$, $p < .001$, $\eta_p^2 = .266$	$F_2(1, 46) = 112.58$, $MSE = 106356$, $p < .001$, $\eta_p^2 = .710$
Group	$F_1(1, 39) = 1.12$, $MSE = 1059$, $p = .297$, $\eta_p^2 = .028$	$F_2(1, 46) = 0.40$, $MSE = 666$, $p = .531$, $\eta_p^2 = .009$
Length	$F_1(1, 39) = 117.70$, $MSE = 431445$, $p < .001$, $\eta_p^2 = .751$	$F_2(1, 46) = 121.69$, $MSE = 257382$, $p < .001$, $\eta_p^2 = .726$
Blocks x Group	$F_1(1, 39) = 9.69$, $MSE = 12177$, $p < .005$, $\eta_p^2 = .199$	$F_2(1, 46) = 14.06$, $MSE = 7430$, $p < .001$, $\eta_p^2 = .234$
Blocks x Length	$F_1(1, 39) = 72.83$, $MSE = 69797$, $p < .001$, $\eta_p^2 = .651$	$F_2(1, 46) = 42.10$, $MSE = 39775$, $p < .001$, $\eta_p^2 = .478$
Group x Length	$F_1(1, 39) = 0.27$, $MSE = 218$, $p = .606$, $\eta_p^2 = .007$	$F_2(1, 46) = 0.10$, $MSE = 159$, $p = .760$, $\eta_p^2 = .002$
Blocks x Group x Length	$F_1(1, 39) = 3.19$, $MSE = 2691$, $p = .082$, $\eta_p^2 = .076$	$F_2(1, 46) = 3.40$, $MSE = 1797$, $p = .072$, $\eta_p^2 = .069$
Between visual and phonological training ($\alpha = 0.01$)		
Block 1 4-letter	$t_1(39) = 0.27$, $p = .786$	$t_2(23) = 0.12$, $p = .908$
Block 1 7-letter	$t_1(39) = 1.58$, $p = .122$	$t_2(23) = 2.00$, $p = .057$
Block 10 4-letter	$t_1(39) = 2.25$, $p = .030$	$t_2(23) = 0.96$, $p = .348$
Block 10 7-letter	$t_1(39) = 3.19$, $p = .003$	$t_2(23) = 2.04$, $p = .053$

4-letter		
Blocks	$F_1(1, 39) = 2.73$, $MSE = 12198$, $p = .106$, $\eta_p^2 = .065$	$F_2(1, 23) = 9.42$, $MSE = 7994$, $p < .005$, $\eta_p^2 = .291$
Group	$F_1(1, 39) = 2.65$, $MSE = 1118$, $p = .112$, $\eta_p^2 = .064$	$F_2(1, 23) = 0.38$, $MSE = 726$, $p = .545$, $\eta_p^2 = .016$
Block x Group	$F_1(1, 39) = 3.14$, $MSE = 1710$, $p = .084$, $\eta_p^2 = .074$	$F_2(1, 46) = 2.12$, $MSE = 938$, $p = .159$, $\eta_p^2 = .085$
7-letter		
Blocks	$F_1(1, 39) = 25.97$, $MSE = 234320$, $p < .001$, $\eta_p^2 = .400$	$F_2(1, 23) = 132.15$, $MSE = 138017$, $p <.001$, $\eta_p^2 = .016$
Group	$F_1(1, 39) = 0.12$, $MSE = 158$, $p = .732$, $\eta_p^2 = .003$	$F_2(1, 23) = 0.06$, $MSE = 88$, $p = .806$, $\eta_p^2 = .003$
Block x Group	$F_1(1, 39) = 8.47$, $MSE = 13159$, $p < .01$, $\eta_p^2 = .178$	$F_2(1, 23) = 13.37$, $MSE = 8288$, $p < .001$, $\eta_p^2 = .368$
Block 1 and 10 comparison (α = 0.01)		
Visual 4-letter	$t_1(39) = 2.14$, $p =.038$	$t_2(23) = 3.53$, $p < .005$
Phonological 4-letter	$t_1(39) = 0.98$, $p = .335$	$t_2(23) = 1.56$, $p = .133$
Visual 7-letter	$t_1(39) = 5.65$, $p < .001$	$t_2(23) = 10.91$, $p < .001$
Phonological 7-letter	$t_1(39) = 3.71$, $p < .001$	$t_2(23) = 7.17$, $p < .001$

2.1.2 Experiment 5

Effects	*By-participants analysis*	*By-items analysis*
NAMING LATENCY (correct, trimmed RTs)		
Overall analysis		
Group	$F_1(1, 46) = 1.05$, $MSE = 71341$, $p = .311$, $\eta_p^2 = .022$	$F_2(1, 22) = 55.79$, $MSE = 35398$, $p < .001$, $\eta_p^2 = .717$
Train	$F_1(1, 46) = 38.43$, $MSE = 33227$, $p < .001$, $\eta_p^2 = .455$	$F_2(1, 22) = 28.31$, $MSE = 16969$, $p < .001$, $\eta_p^2 = .563$
Length	$F_1(1, 46) = 109.73$, $MSE = 332173$, $p < .001$, $\eta_p^2 = .705$	$F_2(1, 22) = 65.12$, $MSE = 175148$, $p < .001$, $\eta_p^2 = .747$
Blocks	$F_1(1, 46) = 11.49$, $MSE = 114333$, $p < .001$, $\eta_p^2 = .200$	$F_2(1, 22) = 144.36$, $MSE = 56204$, $p < .001$, $\eta_p^2 = .868$
Group x Train	$F_1(1, 46) = 0.25$, $MSE = 213$, $p = .622$, $\eta_p^2 = .005$	$F_2(1, 22) = 0.45$, $MSE = 323$, $p = .511$, $\eta_p^2 = .020$
Group x Length	$F_1(1, 46) = 0.10$, $MSE = 315$, $p = .748$, $\eta_p^2 = .002$	$F_2(1, 22) = 0.09$, $MSE = 55$, $p = .771$, $\eta_p^2 = .004$
Group x Blocks	$F_1(1, 46) = 8.68$, $MSE = 86400$, $p < .005$, $\eta_p^2 = .159$	$F_2(1, 22) = 46.56$, $MSE = 43651$, $p < .001$, $\eta_p^2 = .679$
Train x Length	$F_1(1, 46) = 31.05$, $MSE = 23972$, $p < .001$, $\eta_p^2 = .403$	$F_2(1, 22) = 21.50$, $MSE = 12887$, $p < .001$, $\eta_p^2 = .494$
Length x Blocks	$F_1(1, 46) = 7.29$, $MSE = 7597$, $p < .01$, $\eta_p^2 = .137$	$F_2(1, 22) = 9.37$, $MSE = 3649$, $p < .01$, $\eta_p^2 = .299$
Block x Train	$F_1(1, 46) = 17.92$, $MSE = 22357$, $p < .001$, $\eta_p^2 = 280$	$F_2(1, 22) = 6.89$, $MSE = 12049$, $p < .05$, $\eta_p^2 = .239$
Group x Train x Length	$F_1(1, 46) = 4.26$, $MSE = 3290$, $p < .05$, $\eta_p^2 = .085$	$F_2(1, 22) = 1.63$, $MSE = 1175$, $p = .216$, $\eta_p^2 = .069$
Group x Train x Block	$F_1(1, 46) = 0.04$, $MSE = 54$, $p = .836$, $\eta_p^2 = .001$	$F_2(1, 22) = 0.17$, $MSE = 53$, $p = .688$, $\eta_p^2 = .007$
Group x Length x Blocks	$F_1(1, 46) = 1.47$, $MSE = 1528$, $p = .232$, $\eta_p^2 = .031$	$F_2(1, 22) = 0.94$, $MSE = 880$, $p = .343$, $\eta_p^2 = .041$
Train x Length x Blocks	$F_1(1, 46) = 28.19$, $MSE = 21660$, $p < .001$, $\eta_p^2 = .380$	$F_2(1, 22) = 7.37$, $MSE = 12887$, $p < .05$, $\eta_p^2 = .251$

Group x Train x Length x Blocks	$F_1(1, 46) = 1.04$, $MSE = 799$, $p = .313$, $\eta_p^2 = .022$	$F_2(1, 22) = 1.11$, $MSE = 355$, $p = .304$, $\eta_p^2 = .048$
Visual training		
Train	$F_1(1, 23) = 14.39$, $MSE = 14060$, $p < .001$, $\eta_p^2 = .385$	$F_2(1, 22) = 12.07$, $MSE = 6305$, $p < .005$, $\eta_p^2 = .354$
Length	$F_1(1, 23) = 40.53$, $MSE = 156009$, $p < .001$, $\eta_p^2 = .638$	$F_2(1, 22) = 48.52$, $MSE = 84491$, $p < .001$, $\eta_p^2 = .688$
Blocks	$F_1(1, 23) = 31.62$, $MSE = 199757$, $p < .001$, $\eta_p^2 = .579$	$F_2(1, 22) = 128.00$, $MSE = 99459$, $p < .001$, $\eta_p^2 = .853$
Train x Length	$F_1(1, 23) = 33.91$, $MSE = 22512$, $p < .001$, $\eta_p^2 = .596$	$F_2(1, 22) = 20.90$, $MSE = 10923$, $p < .001$, $\eta_p^2 = .487$
Train x Blocks	$F_1(1, 23) = 7.98$, $MSE = 10107$, $p < .01$, $\eta_p^2 = .257$	$F_2(1, 22) = 4.42$, $MSE = 5251$, $p < .05$, $\eta_p^2 = .167$
Length x Blocks	$F_1(1, 23) = 14.37$, $MSE = 7970$, $p < .001$, $\eta_p^2 = .385$	$F_2(1, 22) = 5.22$, $MSE = 4056$, $p < .05$, $\eta_p^2 = .192$
Trainx Length x Blocks	$F_1(1, 23) = 7.53$, $MSE = 7069$, $p < .05$, $\eta_p^2 = .247$	$F_2(1, 22) = 3.77$, $MSE = 4483$, $p = .065$, $\eta_p^2 = .146$
Post-hoc comparison of RTs between Block 1 and 10 using Bonferroni-corrected t-tests ($\propto = .01$)		
Trained 4-letter	$t_1(23) = 3.95$, $p < .001$	$t_2(11) = 5.09$, $p < .001$
Trained 7-letter	$t_1(23) = 8.24$, $p < .001$	$t_2(11) = 8.47$, $p < .001$
Untrained 4-letter	$t_1(23) = 4.84$, $p < .001$	$t_2(11) = 3.76$, $p < .005$
Untrained 7-letter	$t_1(23) = 2.90$, $p < .01$	$t_2(11) = 3.35$, $p < .01$
Compare to Exp 2 in Chapter 2: Post-hoc comparison of RTs between Block 1 and 10 using Bonferroni-corrected t-tests ($\propto = .01$) previous experiment		
Trained 4-letter	$t_1(23) = 5.57$, $p < .001$	$t_2(11) = 8.01$, $p < .001$
Trained 7-letter	$t_1(23) = 6.48$, $p < .001$	$t_2(11) = 12.75$, $p < .001$
Untrained 4-letter	$t_1(23) = 2.94$, $p < .01$	$t_2(11) = 2.80$, $p = .017$

Untrained 7-letter	$t_1(23) = 3.68, p < .001$	$t_2(11) = 3.42, p < .01$
Block 1		
Train	$F_1(1, 23) = 0.13, MSE = 163, p = .726, \eta_p^2 = .005$	$F_2(1, 22) = 0.02, MSE = 24, p = .887, \eta_p^2 = .001$
Length	$F_1(1, 23) = 51.76, MSE = 117250, p < .001, \eta_p^2 = .692$	$F_2(1, 22) = 36.89, MSE = 62785, p < .001, \eta_p^2 = .626$
Train x Length	$F_1(1, 23) = 2.01, MSE = 2176, p = .170, \eta_p^2 = .080$	$F_2(1, 22) = 0.60, MSE = 705, p = .447, \eta_p^2 = .027$
Block 10		
Train	$F_1(1, 23) = 25.34, MSE = 24003, p < .001, \eta_p^2 = .524$	$F_2(1, 22) = 21.52, MSE = 11532, p < .001, \eta_p^2 = .495$
Length	$F_1(1, 23) = 21.85, MSE = 46728, p < .001, \eta_p^2 = .487$	$F_2(1, 22) = 31.55, MSE = 25761, p < .001, \eta_p^2 = .589$
Train x Length	$F_1(1, 23) = 52.72, MSE = 27405, p < .001, \eta_p^2 = .696$	$F_2(1, 22) = 27.44, MSE = 14700, p < .001, \eta_p^2 = .555$
Phonological training		
Train	$F_1(1, 23) = 25.78, MSE = 19380, p < .001, \eta_p^2 = .529$	$F_2(1, 22) = 13.74, MSE = 10987, p < .001, \eta_p^2 = .384$
Length	$F_1(1, 23) = 80.03, MSE = 176479, p < .001, \eta_p^2 = .777$	$F_2(1, 22) = 57.31, MSE = 90713, p < .001, \eta_p^2 = .723$
Blocks	$F_1(1, 23) = .07, MSE = 977, p = .791, \eta_p^2 = .003$	$F_2(1, 22) = 0.72, MSE = 396, p = .405, \eta_p^2 = .032$
Train x Length	$F_1(1, 23) = 5.40, MSE = 4750, p < .05, \eta_p^2 = .190$	$F_2(1, 22) = 3.93, MSE = 3140, p = .06, \eta_p^2 = .151$
Train x Blocks	$F_1(1, 23) = 10.02, MSE = 12304, p < .005, \eta_p^2 = .303$	$F_2(1, 22) = 7.79, MSE = 6851, p < .05, \eta_p^2 = .261$
Length x Blocks	$F_1(1, 23) = 0.76, MSE = 1155, p = .394, \eta_p^2 = .032$	$F_2(1, 22) = 0.86, MSE = 473, p = .364, \eta_p^2 = .038$
Train x Length x Blocks	$F_1(1, 23) = 25.74, MSE = 15390, p < .001, \eta_p^2 = .528$	$F_2(1, 22) = 9.96, MSE = 8759, p < .005, \eta_p^2 = .312$

Block 1		
Train	$F_1(1, 23) = 0.41$, $MSE = 400$, $p = .530$, $\eta_p^2 = .017$	$F_2(1, 22) = 0.32$, $MSE = 243$, $p = .575$, $\eta_p^2 = .015$
Length	$F_1(1, 23) = 45.21$, $MSE = 103097$, $p < .001$, $\eta_p^2 = .663$	$F_2(1, 22) = 41.86$, $MSE = 52140$, $p < .001$, $\eta_p^2 = .655$
Train x Length	$F_1(1, 23) = 2.85$, $MSE = 1520$, $p = .105$, $\eta_p^2 = .110$	$F_2(1, 22) = 0.94$, $MSE = 705$, $p = .343$, $\eta_p^2 = .041$
Block 10		
Train	$F_1(1, 23) = 31.37$, $MSE = 31284$, $p < .001$, $\eta_p^2 = .577$	$F_2(1, 22) = 18.93$, $MSE = 17595$, $p < .001$, $\eta_p^2 = .462$
Length	$F_1(1, 23) = 51.22$, $MSE = 74538$, $p < .001$, $\eta_p^2 = .690$	$F_2(1, 22) = 44.02$, $MSE = 39045$, $p < .001$, $\eta_p^2 = .667$
Train x Length	$F_1(1, 23) = 19.70$, $MSE = 18621$, $p < .001$, $\eta_p^2 = .461$	$F_2(1, 22) = 12.04$, $MSE = 11194$, $p < .005$, $\eta_p^2 = .354$
Post-hoc comparison of RTs between Block 1 and 10 using Bonferroni-corrected t-tests (∝ = .01)		
Trained 4-letter	$t_1(23) = 0.14$, $p = .888$	$t_2(11) = 0.22$, $p = .833$
Trained 7-letter	$t_1(23) = 1.79$, $p = .087$	$t_2(11) = 4.37$, $p < .001$
Untrained 4-letter	$t_1(23) = 0.10$, $p = .918$	$t_2(11) = 0.18$, $p = .861$
Untrained 7-letter	$t_1(23) = 1.30$, $p = .208$	$t_2(11) = 2.47$, $p < .05$

2.1.3 Experiment 6

Effects	***By-participants analysis***	***By-items analysis***
Red dots catch trials		
Accuracy	$H(104) = 5.38$, $p = .146$	$W(6) = 0.20$, $p = .319$
Reaction RTs	$F_1(3, 100) = 1.73$, $MSE = 2363$, $p = .165$, $\eta_p^2 = .049$	$F_2(1.95, 9.77) = 13.27$, $MSE = 862$, $p < .005$, $\eta_p^2 = .726$
NAMING LATENCY (correct, trimmed RTs)		
Overall analysis(all 4 conditions)		
Training	$F_1(1, 100) = 98.45$, $MSE = 82044$, $p < .001$, $\eta_p^2 = .496$	$F_2(1, 29) = 42.46$, $MSE = 97454$, $p < .001$, $\eta_p^2 = .594$
Conditions	$F_1(3, 100) = 0.75$, $MSE = 10079$, $p = .526$, $\eta_p^2 = .022$	$F_2(3, 87) = 12.52$, $MSE = 11250$, $p < .001$, $\eta_p^2 = .302$
Training x Conditions	$F_1(3, 100) = 3.46$, $MSE = 2882$, $p < .05$, $\eta_p^2 = .094$	$F_2(1.50, 43.39) = 1.35$, $MSE = 7164$, $p = .264$, $\eta_p^2 = .045$
Post-hoc comparison of RTs between trained and untrained items with Bonferroni correction ($\propto = .01$)		
Read aloud	$t_1(25) = 5.67$, $p < .001$	$t_2(29) = 6.19$, $p < .001$
Hear-and-repeat	$t_1(25) = 6.85$, $p < .001$	$t_2(29) = 3.40$, $p < .005$
Hear-and-repeat with orthographic distractor	$t_1(25) = 4.81$, $p < .001$	$t_2(29) = 2.81$, $p = .009$
Hear-and-repeat with non-orthographic distractor	$t_1(25) = 3.48$, $p < .005$	$t_2(29) = 1.65$, $p = .109$
Non-parametric analysis for difference score		
Difference between trained and untrained items(all 4 conditions)	*Kruskal-Wallis Test*	*Friedman's Test*
Conditions	$H(3) = 8.39$, $p < .05$	$X^2(3) = 4.36$, $p = .225$

Post-hoc comparison of difference RTs between each condition using Bonferroni-correction (∝ = .017)	***Mann-Whitney U Test***	**Wilcoxon signed rank test**
Read aloud and Hear-and-repeat	*U*(50) = 11.56, *Z* = 1.38, *p* = .167	*W*(29) = 134.00, *Z* = 2.03, *p* = .043
Read aloud and Hear-and-repeat with orthographic distractor	*U*(50) = 20.52, *Z* = 2.45, *p* = .014	*W*(29) = 135.00, *Z* = 2.01, *p* = .045
Read aloud and Hear-and-repeat with non-orthographic distractor	*U*(50) = 21.15, *Z* = 2.53, *p* = .011	*W*(29) = 165.00, *Z* = 1.39, *p* = .165

3 Appendix 3

3.1 Analysis of the result in Experiment 7 of Chapter 4.

3.1.1 Experiment 7

Effects	***By-participants analysis***	***By-items analysis***
NAMING LATENCIES (RTs)		
Day 1, block 1: high-frequency words, low-frequency words and nonwords		
Stimulus type	$F_1(1.36, 32.67) = 26.13, MSE = 96738, p < .001, \eta_p^2 = .521$	$F_2(2, 66) = 52.29, MSE = 32323, p < .001, \eta_p^2 = .613$
Length	$F_1 1, 24) = 23.14, MSE = 48955, p < .001, \eta_p^2 = .491$	$F_2(1, 66) = 37.96, MSE = 23465, p < .001, \eta_p^2 = .365$
Stimulus type x Length	$F_1(2, 48) = 40.15, MSE = 21168, p < .001, \eta_p^2 = .626$	$F_2(2, 66) = 16.37, MSE = 10120, p < .001, \eta_p^2 = .332$
Day 1, block 1: high-frequency words *vs.* low-frequency words		
Stimulus type	$F_1(1, 24) = 5.89, MSE = 5340, p < .05, \eta_p^2 = .197$	$F_2(1, 44) = 4.05, MSE = 2523, p < .05, \eta_p^2 = .084$
Length	$F_1(1, 24) = 4.71, MSE = 3933, p < .05, \eta_p^2 = .164$	$F_2(1, 44) = 3.03, MSE = 1885, p = .089, \eta_p^2 = .064$
Stimulus type x Length	$F_1(1, 24) = 1.42, MSE = 603, p = .245, \eta_p^2 = .056$	$F_2(1, 44) = 0.38, MSE = 237, p = .540, \eta_p^2 = .009$
Day 1, block 1: high-frequency words *vs.* nonwords		
Stimulus type	$F_1(1, 24) = 30.75, MSE = 118594, p < .001, \eta_p^2 = .562$	$F_2(1, 44) = 96.69, MSE = 58064, p < .001, \eta_p^2 = .687$
Length	$F_1(1, 24) = 31.70, MSE = 51689, p < .001, \eta_p^2 = .569$	$F_2(1, 44) = 41.68, MSE = 25028, p < .001, \eta_p^2 = .486$
Stimulus type x Length	$F_1(1, 24) = 54.77, MSE = 35796, p < .001, \eta_p^2 = .695$	$F_2(1, 44) = 28.22, MSE = 16948 p < .001, \eta_p^2 = .391$
Day 1, block 1: low-frequency words *vs.* nonwords		
Stimulus type	$F_1(1, 24) = 26.33, MSE = 73605, p < .001, \eta_p^2 = .523$	$F_2(1, 44) = 57.60, MSE = 36381, p < .001, \eta_p^2 = .567$
Length	$F_1(1, 24) = 26.68, MSE = 63458, p < .001, \eta_p^2 = .536$	$F_2(1, 44) = 47.72, MSE = 30138, p < .001, \eta_p^2 = .520$
Stimulus type x Length	$F_1(1, 24) = 53.82, MSE = 27106, p < .001, \eta_p^2 = .692$	$F_2(1, 44) = 20.86, MSE = 13175, p < .001, \eta_p^2 = .322$

Bonferroni corrected t-tests (α = .017): effects of length on high-frequency words, low-frequency words and nonwords in day 1, block 1		
High frequency words: 4 vs. 7 letters	$t_1(24) = 1.40$, $p = .176$	$t_2(22) = 0.82$, $p = .424$
Low frequency words: 4 vs. 7 letters	$t_1(24) = 2.07$, $p = .049$	$t_2(22) = 1.63$, $p = .118$
Nonwords: 4 vs. 7 letters	$t_1(24) = 6.74$, $p < .001$	$t_2(22) = 8.26$, $p < .001$
Day 1, blocks 1-10: high-frequency words, low-frequency words and nonwords		
Blocks	$F_1(2.89, 69.34) = 21.81$, $MSE = 182541$, $p < .001$, $\eta_p^2 = .476$	$F_2(9, 594) = 110.09$, $MSE = 28199$, $p < .001$, $\eta_p^2 = .625$
Stimulus type	$F_1(1.12, 26.81) = 24.52$, $MSE = 210576$, $p < .001$, $\eta_p^2 = .505$	$F_2(2, 66) = 17.99$, $MSE = 56926$, $p < .001$, $\eta_p^2 = .353$
Length	$F_1(1, 24) = 5.82$, $MSE = 40577$, $p < .05$, $\eta_p^2 = .195$	$F_2(1, 66) = 6.15$, $MSE = 19465$, $p < .05$, $\eta_p^2 = .085$
Blocks x Stimulus type	$F_1(6.44, 154.59) = 6.38$, $MSE = 9169$, $p < .001$, $\eta_p^2 = .210$	$F_2(18, 594) = 6.36$, $MSE = 1629$, $p < .001$, $\eta_p^2 = .162$
Blocks x Length	$F_1(4.50, 107.98) = 8.32$, $MSE = 7140$, $p < .001$, $\eta_p^2 = .257$	$F_2(9, 594) = 6.75$, $MSE = 1728$, $p < .001$, $\eta_p^2 = .093$
Stimulus type x Length	$F_1(1.17, 27.97) = 22.20$, $MSE = 66816$, $p < .001$, $\eta_p^2 = .481$	$F_2(2, 66) = 5.98$, $MSE = 18939$, $p < .005$, $\eta_p^2 = .154$
Blocks x Stimulus type x Length	$F_1(7.28, 174.70) = 3.63$, $MSE = 3277$, $p < .01$, $\eta_p^2 = .131$	$F_2(18, 594) = 2.45$, $MSE = 628$, $p < .01$, $\eta_p^2 = .069$
Day 1, blocks 1-10: high-frequency words vs. low-frequency words		
Blocks	$F_1(3.55, 85.10) = 13.14$, $MSE = 61313$, $p < .001$, $\eta_p^2 = .354$	$F_2(9, 396) = 45.95$, $MSE = 11546$, $p < .001$, $\eta_p^2 = .511$
Stimulus type	$F_1(1, 24) = 15.63$, $MSE = 13981$, $p < .001$, $\eta_p^2 = .394$	$F_2(1, 44) = 2.06$, $MSE = 6670$, $p = .159$, $\eta_p^2 = .045$
Length	$F_1(1, 24) = 0.01$, $MSE = 17$, $p = .948$, $\eta_p^2 = .000$	$F_2(1, 44) = 0.01$, $MSE = 2$, $p = .978$, $\eta_p^2 = .000$
Blocks x Stimulus type	$F_1(4.80, 115.08) = 0.91$, $MSE = 687$, $p = .474$, $\eta_p^2 = .037$	$F_2(9, 396) = 0.73$, $MSE = 183$, $p = .682$, $\eta_p^2 = .016$
Blocks x Length	$F_1(5.93, 142.22) = 1.98$, $MSE = 921$, $p = .073$, $\eta_p^2 = .076$	$F_2(9, 396) = 1.19$, $MSE = 298$, $p = .303$, $\eta_p^2 = .026$
Stimulus type x Length	$F_1(1, 24) = 0.16$, $MSE = 48$, $p = .689$, $\eta_p^2 = .007$	$F_2(1, 44) = 0.01$, $MSE = 16$, $p = .944$, $\eta_p^2 = .000$
Blocks x Stimulus type x Length	$F_1(9, 216) = 0.74$, $MSE = 234$, $p = .670$, $\eta_p^2 = .030$	$F_2(9, 396) = 0.42$, $MSE = 104$, $p = .927$, $\eta_p^2 = .009$

Day 1, blocks 1-10: high-frequency words vs. nonwords		
Blocks	$F_1(3.06, 73.51) = 23.37$, $MSE = 135351$, $p < .001$, $\eta_p^2 = .493$	$F_2(9, 396) = 90.01$, $MSE = 22233$, $p < .001$, $\eta_p^2 = .672$
Stimulus type	$F_1(1, 24) = 26.42$, $MSE = 217621$, $p < .001$, $\eta_p^2 = .524$	$F_2(1, 44) = 32.83$, $MSE = 105335$, $p < .001$, $\eta_p^2 = .427$
Length	$F_1(1, 24) = 8.41$, $MSE = 58345$, $p < .01$, $\eta_p^2 = .259$	$F_2(1, 44) = 8.77$, $MSE = 28127$, $p < .005$, $\eta_p^2 = .166$
Blocks x Stimulus type	$F_1(4.33, 103.91) = 9.45$, $MSE = 11074$, $p < .001$, $\eta_p^2 = .282$	$F_2(9, 396) = 10.69$, $MSE = 2640$, $p < .001$, $\eta_p^2 = .195$
Blocks x Length	$F_1(4.88, 117.07) = 7.50$, $MSE = 6645$, $p < .001$, $\eta_p^2 = .238$	$F_2(9, 396) = 7.14$, $MSE = 1764$, $p < .001$, $\eta_p^2 = .140$
Stimulus type x Length	$F_1(1, 24) = 26.48$, $MSE = 60046$, $p < .001$, $\eta_p^2 = .525$	$F_2(1, 44) = 9.02$, $MSE = 28948$, $p < .005$, $\eta_p^2 = .170$
Blocks x Stimulus type x Length	$F_1(9, 216) = 5.48$, $MSE = 2064$, $p < .001$, $\eta_p^2 = .186$	$F_2(9, 396) = 3.91$, $MSE = 966$, $p < .001$, $\eta_p^2 = .082$
Day 1, blocks 1-10: low-frequency words vs. nonwords		
Blocks	$F_1(2.68, 64.27) = 24.18$, $MSE = 168915$, $p < .001$, $\eta_p^2 = .502$	$F_2(9, 396) = 89.90$, $MSE = 24288$, $p < .001$, $\eta_p^2 = .671$
Stimulus type	$F_1(1, 24) = 23.06$, $MSE = 121283$, $p < .001$, $\eta_p^2 = .490$	$F_2(1, 44) = 19.36$, $MSE = 58992$, $p < .001$, $\eta_p^2 = .306$
Length	$F_1(1, 24) = 10.10$, $MSE = 61731$, $p < .005$, $\eta_p^2 = .296$	$F_2(1, 44) = 9.68$, $MSE = 29487$, $p < .005$, $\eta_p^2 = .180$
Blocks x Stimulus type	$F_1(3.90, 93.63) = 7.19$, $MSE = 9574$, $p < .001$, $\eta_p^2 = .230$	$F_2(9, 396) = 7.69$, $MSE = 2078$, $p < .001$, $\eta_p^2 = .149$
Blocks x Length	$F_1(4.21, 100.10) = 9.74$, $MSE = 9103$, $p < .001$, $\eta_p^2 = .289$	$F_2(9, 396) = 7.51$, $MSE = 2029$, $p < .001$, $\eta_p^2 = .146$
Stimulus type x Length	$F_1(1, 24) = 20.99$, $MSE = 56708$, $p < .001$, $\eta_p^2 = .466$	$F_2(1, 44) = 9.06$, $MSE = 27601$, $p < .005$, $\eta_p^2 = .171$
Blocks x Stimulus type x Length	$F_1(5.27, 126.44) = 4.15$, $MSE = 2868$, $p < .001$, $\eta_p^2 = .147$	$F_2(9, 396) = 3.03$, $MSE = 819$, $p < .005$, $\eta_p^2 = .064$
Day 1, blocks 1-10: high-frequency words only		
Blocks	$F_1(4.48, 107.52) = 10.15$, $MSE = 21601$, $p < .001$, $\eta_p^2 = .297$	$F_2(9, 198) = 22.44$, $MSE = 5118$, $p < .001$, $\eta_p^2 = .505$
Length	$F_1(1, 24) = 0.01$, $MSE = 6$, $p = .952$, $\eta_p^2 = .000$	$F_2(1, 22) = 0.01$, $MSE = 3$, $p = .977$, $\eta_p^2 = .000$
Blocks x Length	$F_1(9, 216) = 0.90$, $MSE = 285$, $p = .531$, $\eta_p^2 = .036$	$F_2(9, 198) = 0.62$, $MSE = 142$, $p = .778$, $\eta_p^2 = .027$

Day 1, blocks 1-10: low-frequency words only		
Blocks	$F_1(3.44, 82.48) = 11.67, MSE = 36064, p < .001, \eta_p^2 = .327$	$F_2(9, 198) = 24.09, MSE = 6611, p < .001, \eta_p^2 = .523$
Length	$F_1(1, 24) = 0.04, MSE = 53, p = .839, \eta_p^2 = .002$	$F_2(1, 22) = 0.01, MSE = 16, p = .944, \eta_p^2 = .000$
Blocks x Length	$F_1(5.05, 121.24) = 1.83, MSE = 989, p = .111, \eta_p^2 = .071$	$F_2(9, 198) = 0.95, MSE = 260, p = .484, \eta_p^2 = .041$
Day 1, blocks 1-10: nonwords only		
Blocks	$F_1(2.76, 66.19) = 27.54, MSE = 132616, p < .001, \eta_p^2 = .534$	$F_2(9, 198) = 74.30, MSE = 19755, p < .001, \eta_p^2 = .772$
Length	$F_1(1, 24) = 15.68, MSE = 118385, p < .01, \eta_p^2 = .395$	$F_2(1, 22) = 18.95, MSE = 57073, p < 001, \eta_p^2 = .463$
Blocks x Length	$F_1(4.48, 107.42) = 9.98, MSE = 10818, p < .001, \eta_p^2 = .294$	$F_2(9, 198) = 9.73, MSE = 2588, p < .001\ \eta_p^2 = .307$
Bonferroni corrected t-tests (α = .005): effects of length on RTs to nonwords in each of blocks 1-10 on day 1		
Block 1	$t_1(24) = 6.74, p < .001$	$t_2(22) = 8.26, p < .001$
Block 2	$t_1(24) = 3.10, p = .005$	$t_2(22) = 4.49\ p < .001$
Block 3	$t_1(24) = 3.09, p = .005$	$t_2(22) = 3.38, p = .003$
Block 4	$t_1(24) = 3.64, p < .001$	$t_2(22) = 2.43, p = .025$
Block 5	$t_1(24) = 2.68, p = .013$	$t_2(22) = 3.10, p = .005$
Block 6	$t_1(24) = 3.48, p = .002$	$t_2(22) = 2.67, p = .014$
Block 7	$t_1(24) = 0.69, p = .494$	$t_2(22) = 0.59, p =.559$
Block 8	$t_1(24) = 3.54, p = .002$	$t_2(22) = 2.46, p = .023$
Block 9	$t_1(24) = 1.61, p = .121$	$t_2(22) = 2.27, p = .033$
Block 10	$t_1(24) = 2.54, p = .018$	$t_2(22) = 2.38, p =.027$
Retention across 28 days: day 28 block 1 *vs*. day 1 block 10		
Delay	$F_1(1, 24) = 8.94, MSE = 44176, p < .01, \eta_p^2 = .271$	$F_2(1, 66) = 76.59, MSE = 20897, p < .001, \eta_p^2 = .537$
Stimulus type	$F_1(1.39, 33.30) = 16.59, MSE = 19209, p < .001, \eta_p^2 = .409$	$F_2(2, 66) = 8.78, MSE = 6217, p < .001, \eta_p^2 = .210$
Length	$F_1(1, 24) = 4.34, MSE = 7588, p < .05, \eta_p^2 = .153$	$F_2(1, 66) = 4.96, MSE = 3515, p < .05, \eta_p^2 = .070$
Delay x Stimulus type	$F_1(2, 48) = 6.75, MSE = 3415, p < .005, \eta_p^2 = .220$	$F_2(2, 66) = 5.69, MSE = 1553, p < .01, \eta_p^2 = .147$
Delay x Length	$F_1(1, 24) = 2.84, MSE = 1599, p = .105, \eta_p^2 = .106$	$F_2(1, 66) = 2.60, MSE = 708, p = .112, \eta_p^2 = .038$
Stimulus type x Length	$F_1(1.57, 37.57) = 18.91, MSE = 14398, p < .001, \eta_p^2 = .441$	$F_2(2, 66) = 7.42, MSE = 5253, p < .01, \eta_p^2 = .184$
Delay x Stimulus type x Length	$F_1(2, 48) = 2.56, MSE = 1076, p = .088, \eta_p^2 = .096$	$F_2(2, 66) = 1.71, MSE = 467, p = .189, \eta_p^2 = .049$

Bonferroni corrected t-tests (α = .017): day 28 block 1 vs. day 1 block 10 for high frequency words, low frequency words and nonwords (4 and 7 letters combined)		
High frequency words	$t_1(24) = 2.29, p = .031$	$t_2(23) = 5.72, p < .001$
Low frequency words	$t_1(24) = 1.97, p = .060$	$t_2(23) = 2.50, p = .020$
Nonwords	$t_1(24) = 3.79, p < .001$	$t_2(23) = 7.20, p < .001$
Day 28, blocks 1-10: high-frequency words, low-frequency words and nonwords		
Stimulus type	$F_1(2, 48) = 35.89, MSE = 34711, p < .001, \eta_p^2 = .599$	$F_2(2, 66) = 7.57, MSE = 16514, p < .01, \eta_p^2 = .187$
Blocks	$F_1(3.18, 76.44) = 2.10, MSE = 15720, p = .103, \eta_p^2 = .080$	$F_2(9, 594) = 9.50, MSE = 2620, p < .001, \eta_p^2 = .126$
Length	$F_1(1, 24) = 0.37, MSE = 1357, p = .549, \eta_p^2 = .015$	$F_2(1, 66) = 0.28, MSE = 619, p = .596, \eta_p^2 = .004$
Blocks x Stimulus type	$F_1(8.54, 204.87) = 2.81, MSE = 2495, p < .01, \eta_p^2 = .105$	$F_2(18, 594) = 2.02, MSE = 556, p < .01, \eta_p^2 = .058$
Stimulus type x Length	$F_1(2, 48) = 12.05, MSE = 7554, p < .001, \eta_p^2 = .334$	$F_2(2, 66) = 1.64, MSE = 3577, p = .202, \eta_p^2 = .047$
Blocks x Length	$F_1(5.17, 124.13) = 3.49, MSE = 2389, p < .01, \eta_p^2 = .127$	$F_2(9, 594) = 2.33, MSE = 644, p < .05, \eta_p^2 = .034$
Blocks x Stimulus type x Length	$F_1(8.74, 209.65) = 2.03, MSE = 1912, p < .05, \eta_p^2 = .078$	$F_2(18, 594) = 1.56, MSE = 429, p = .066, \eta_p^2 = .045$
Day 28, blocks 1-10: high-frequency words vs. low-frequency words		
Blocks	$F_1(3.77, 90.54) = 1.54, MSE = 5088, p = .201, \eta_p^2 = .060$	$F_2(9, 396) = 3.79, MSE = 974, p < .001, \eta_p^2 = .079$
Stimulus type	$F_1(1, 24) = 0.42, MSE = 174, p = .522, \eta_p^2 = .017$	$F_2(1, 44) = 1.87, MSE = 4034, p = .179, \eta_p^2 = .041$
Length	$F_1(1, 24) = 1.66, MSE = 966, p = .210, \eta_p^2 = .065$	$F_2(1, 44) = 0.37, MSE = 809, p = .544, \eta_p^2 = .008$
Blocks x Stimulus type	$F_1(3.58, 85.94) = 0.56, MSE = 1359, p = .674, \eta_p^2 = .023$	$F_2(9, 396) = 1.96, MSE = 504, p < .05, \eta_p^2 = .043$
Blocks x Length	$F_1(5.45, 130.78) = 1.92, MSE = 1387, p = .089, \eta_p^2 = .074$	$F_2(9, 396) = 0.76, MSE = 195, p = .657, \eta_p^2 = .017$
Stimulus type x Length	$F_1(1, 24) = 0.10, MSE = 34, p = .754, \eta_p^2 = .004$	$F_2(1, 44) = 0.01, MSE = 22, p = .919, \eta_p^2 = .000$
Blocks x Stimulus type x Length	$F_1(9, 216) = 2.77, MSE = 1309, p < .005, \eta_p^2 = .103$	$F_2(9, 396) = 0.63, MSE = 162, p = .772, \eta_p^2 = .014$
Day 28, blocks 1-10: high-frequency words vs. nonwords		
Blocks	$F_1(3.32,79.71) = 6.82, MSE = 25012, p < .001, \eta_p^2 = .221$	$F_2(9, 396) = 9.43, MSE = 2511, p < .001, \eta_p^2 = .176$
Stimulus type	$F_1(1, 24) = 1.87, MSE = 1417, p = .185, \eta_p^2 = .072$	$F_2(1, 44) = 15.21, MSE = 32117, p < .001, \eta_p^2 = .257$
Length	$F_1(1, 24) = 7.02, MSE = 5661, p < .05, \eta_p^2 = .226$	$F_2(1, 44) = 1.03, MSE = 2173, p = .316, \eta_p^2 = .023$
Blocks x Stimulus type	$F_1(3.47, 83.37) = 3.81, MSE = 10858, p < .01, \eta_p^2 = .137$	$F_2(9, 396) = 1.69, MSE = 450, p = .090, \eta_p^2 = .037$
Blocks x Length	$F_1(5.87, 140.94) = 3.99, MSE = 3338, p < .001, \eta_p^2 = .142$	$F_2(9, 396) = 2.54, MSE = 675, p < .01, \eta_p^2 = .054$

Stimulus type x Length	$F_1(1, 24) = 1.03$, $MSE = 411$, $p = .319$, $\eta_p^2 = .041$	$F_2(1, 44) = 2.34$, $MSE = 4945$, $p = .133$, $\eta_p^2 = .051$
Blocks x Stimulus type x Length	$F_1(9, 216) = 3.27$, $MSE = 1851$, $p < .001$, $\eta_p^2 = .120$	$F_2(9, 396) = 2.24$, $MSE = 598$, $p < .05$, $\eta_p^2 = .049$
Day 28, blocks 1-10: low-frequency words vs. nonwords		
Blocks	$F_1(3.52, 84.50) = 3.54$, $MSE = 13668$, $p < .05$, $\eta_p^2 = .129$	$F_2(9, 396) = 7.63$, $MSE = 2306$, $p < .001$, $\eta_p^2 = .148$
Stimulus type	$F_1(1, 24) = 0.80$, $MSE = 521$, $p = .379$, $\eta_p^2 = .032$	$F_2(1, 44) = 5.87$, $MSE = 13386$, $p < .05$, $\eta_p^2 = .118$
Length	$F_1(1, 24) = 3.45$, $MSE = 2334$, $p = .075$, $\eta_p^2 = .126$	$F_2(1, 44) = 0.77$, $MSE = 1755$, $p = .385$, $\eta_p^2 = .017$
Blocks x Stimulus type	$F_1(4.65, 111.61) = 4.20$, $MSE = 8012$, $p < .005$, $\eta_p^2 = .149$	$F_2(9, 396) = 2.36$, $MSE = 714$, $p < .05$, $\eta_p^2 = .051$
Blocks x Length	$F_1(5.20, 124.77) = 4.88$, $MSE = 4653$, $p < .001$, $\eta_p^2 = .169$	$F_2(9, 396) = 2.80$, $MSE = 846$, $p < .005$, $\eta_p^2 = .060$
Stimulus type x Length	$F_1(1, 24) = 8.53$, $MSE = 3333$, $p < .01$, $\eta_p^2 = .262$	$F_2(1, 44) = 2.47$, $MSE = 5633$, $p = .123$, $\eta_p^2 = .053$
Blocks x Stimulus type x Length	$F_1(4.43, 106.29) = 2.51$, $MSE = 2943$, $p < .05$, $\eta_p^2 = .095$	$F_2(9, 396) = 1.74$, $MSE = 528$, $p = .078$, $\eta_p^2 = .038$
Day 28: high-frequency words only		
Blocks	$F_1(3.72, 89.17) = 1.38$, $MSE = 3590$, $p = .251$, $\eta_p^2 = .054$	$F_2(9, 198) = 3.21$, $MSE = 710$, $p < .01$, $\eta_p^2 = .127$
Length	$F_1(1, 24) = 0.39$, $MSE = 567$, $p = .537$, $\eta_p^2 = .016$	$F_2(1, 22) = 0.14$, $MSE = 281$, $p = .711$, $\eta_p^2 = .006$
Blocks x Length	$F_1(9, 216) = 0.70$, $MSE = 263$, $p = .712$, $\eta_p^2 = .028$	$F_2(9, 198) = 0.58$, $MSE = 128$, $p = .812$, $\eta_p^2 = .026$
Day 28: low-frequency words only		
Blocks	$F_1(3.46, 82.98) = 1.55$, $MSE = 4170$, $p = .203$, $\eta_p^2 = .061$	$F_2(9, 198) = 2.62$, $MSE = 769$, $p < .01$, $\eta_p^2 = .106$
Length	$F_1(1, 24) = 0.60$, $MSE = 1146$, $p = .447$, $\eta_p^2 = .024$	$F_2(1, 22) = 0.24$, $MSE = 550$, $p = .632$, $\eta_p^2 = .011$
Blocks x Length	$F_1(9, 216) = 1.10$, $MSE = 475$, $p = .366$, $\eta_p^2 = .044$	$F_2(9, 198) = 0.78$, $MSE = 228$, $p = .637$, $\eta_p^2 = .034$
Day 28: nonwords only		
Blocks	$F_1(4.06, 97.32) = 3.52$, $MSE = 10755$, $p < .05$, $\eta_p^2 = .128$	$F_2(5.00, 110.14) = 7.23$, $MSE = 4047$, $p < .001$, $\eta_p^2 = .247$
Length	$F_1(1, 24) = 9.40$, $MSE = 14754$, $p < .01$, $\eta_p^2 = .281$	$F_2(1, 22) = 3.07$, $MSE = 6837$, $p = .094$, $\eta_p^2 = .122$
Blocks x Length	$F_1(4.97, 119.35) = 5.01$, $MSE = 4506$, $p < .001$, $\eta_p^2 = .173$	$F_2(9, 198) = 3.68$, $MSE = 1145$, $p < .001$, $\eta_p^2 = .143$
Bonferroni corrected t-tests ($\alpha = .005$): effects of length on RTs to nonwords in each of blocks 1-10 on day 28		
Block 1	$t_1(24) = 4.38$, $p < .001$	$t_2(22) = 4.89$, $p < .001$
Block 2	$t_1(24) = 1.83$, $p = .080$	$t_2(22) = 1.78$, $p = .089$

Block 3	$t_1(24) = 1.63$, $p = .116$	$t_2(22) = 1.24$, $p = .230$
Block 4	$t_1(24) = 1.17$, $p = .254$	$t_2(22) = 0.05$, $p = .584$
Block 5	$t_1(24) = 2.35$, $p = .027$	$t_2(22) = 2.23$, $p = .037$
Block 6	$t_1(24) = 0.25$, $p = .803$	$t_2(22) = 0.21$, $p = .839$
Block 7	$t_1(24) = 0.04$, $p = .972$	$t_2(22) = 0.03$, $p = .975$
Block 8	$t_1(24) = 2.56$, $p = .017$	$t_2(22) = 1.37$, $p =.184$
Block 9	$t_1(24) = 0.45$, $p = .657$	$t_2(22) = 0.32$, $p = .751$
Block 10	$t_1(24) = 1.01$, $p = .323$	$t_2(22) = 0.62$, $p = .542$
Overall analysis of days 1 and 28		
Day	$F_1(1, 24) = 2.14$, $MSE = 77563$, $p = .156$, $\eta_p^2 = .082$	$F_2(1, 66) = 92.88$, $MSE = 37402$, $p < .001$, $\eta_p^2 = .585$
Blocks	$F_1(3.05, 73.27) = 13.22$, $MSE = 137272$, $p < .001$, $\eta_p^2 = .355$	$F_2(9, 594) = 61.17$, $MSE = 22294$, $p < .001$, $\eta_p^2 = .481$
Stimulus type	$F_1(1.18, 28.38) = 37.89$, $MSE = 236226$, $p < .001$, $\eta_p^2 = .612$	$F_2(2, 66) = 13.59$, $MSE = 67187$, $p < .001$, $\eta_p^2 = .292$
Length	$F_1(1, 24) = 3.03$, $MSE = 28388$, $p = .095$, $\eta_p^2 = .112$	$F_2(1, 66) = 2.73$, $MSE = 13515$, $p = .103$, $\eta_p^2 = .040$
Day x Blocks	$F_1(3.09, 74.23) = 9.71$, $MSE = 51211$, $p < .001$, $\eta_p^2 = .288$	$F_2(9, 594) = 50.90$, $MSE = 8525$, $p < .001$, $\eta_p^2 = .435$
Day x Stimulus type	$F_1(1.33, 31.91) = 6.09$, $MSE = 19044$, $p < .05$, $\eta_p^2 = .202$	$F_2(2, 66) = 15.53$, $MSE = 6253$, $p < .001$, $\eta_p^2 = .320$
Day x Length	$F_1(1, 24) = 10.72$, $MSE = 13546$, $p < .01$, $\eta_p^2 = .309$	$F_2(1, 66) = 16.31$, $MSE = 6570$, $p < .001$, $\eta_p^2 = .198$
Blocks x Stimulus type	$F_1(6.98, 167.47) = 6.46$, $MSE = 8942$, $p < .001$, $\eta_p^2 = .212$	$F_2(18, 594) = 4.60$, $MSE = 1675$, $p < .001$, $\eta_p^2 = .122$
Blocks x Length	$F_1(4.60, 110.37) = 9.86$, $MSE = 8782$, $p < .001$, $\eta_p^2 = .291$	$F_2(9, 594) = 5.88$, $MSE = 2143$, $p < .001$, $\eta_p^2 = .082$
Stimulus type x Length	$F_1(1.30, 31.29) = 33.01$, $MSE = 61872$, $p < .001$, $\eta_p^2 = .579$	$F_2(2, 66) = 3.94$, $MSE = 19464$, $p < .05$, $\eta_p^2 = .107$
Day x Blocks x Stimulus type	$F_1(7.42, 177.99) = 2.51$, $MSE = 2422$, $p < .05$, $\eta_p^2 = .095$	$F_2(18, 594) = 3.05$, $MSE = 510$, $p < .001$, $\eta_p^2 = .085$
Day x Blocks x Length	$F_1(5.63, 135.05) = 1.24$, $MSE = 721$, $p = .293$, $\eta_p^2 = .049$	$F_2(9, 594) = 1.37$, $MSE = 229$, $p = .199$, $\eta_p^2 = .020$
Day x Stimulus type x Length	$F_1(1.50, 35.89) = 5.31$, $MSE = 8228$, $p < .05$, $\eta_p^2 = .181$	$F_2(2, 66) = 7.58$, $MSE = 3051$, $p < .001$, $\eta_p^2 = .187$
Blocks x Stimulus type x Length	$F_1(8.17, 196.10) = 4.39$, $MSE = 4339$, $p < .001$, $\eta_p^2 = .155$	$F_2(18, 594) = 2.52$, $MSE = 917$, $p < .01$, $\eta_p^2 = .071$
Day x Blocks x Stimulus type x Length	$F_1(7.92, 190.11) = 0.76$, $MSE = 644$, $p = .639$, $\eta_p^2 = .031$	$F_2(18, 594) = 0.83$, $MSE = 140$, $p = .661$, $\eta_p^2 = .025$

4 Appendix 4

4.1 Analysis of the result in Experiment 8 of Chapter 5.

4.1.1 Experiment 8

Effects	***By-participants analysis***	***By-items analysis***
NAMING LATENCY (correct, trimmed RTs)		
Overall analysis		
Group	$F_1(1, 58) = 25.81$, $MSE = 9291901$, $p < .001$, $\eta_p^2 = .308$	$F_2(1, 22) = 1839.78$, $MSE = 3725512$, $p < .001$, $\eta_p^2 = .988$
Day	$F_1(1, 58) = 71.39$, $MSE = 3216320$, $p < .001$, $\eta_p^2 = .552$	$F_2(1, 22) = 802$, $MSE = 1300873$, $p < .001$, $\eta_p^2 = .973$
Length	$F_1(1, 58) = 84.80$, $MSE = 572957$, $p < .001$, $\eta_p^2 = .594$	$F_2(1, 22) = 17.93$, $MSE = 233620$, $p < .001$, $\eta_p^2 = .449$
Blocks	F_1 $(2.67, 154.61) = 63.56$, $MSE = 1389879$, $p < .001$, $\eta_p^2 = .523$	$F_2(5.54, 121.97) = 98.73$, $MSE = 268546$, $p < .001$, $\eta_p^2 = .818$
Day x Group	F_1 $(1, 58) = 10.54$, $MSE = 474883$, $p < .005$, $\eta_p^2 = .154$	$F_2(1, 22) = 211.26$, $MSE = 190801$, $p < .001$, $\eta_p^2 = .906$
Day x Length	F_1 $(1, 58) = 83.35$, $MSE = 130627$, $p < .001$, $\eta_p^2 = .590$	$F_2(1, 22) = 32.66$, $MSE = 53002$, $p < .001$, $\eta_p^2 = .598$
Group x Length	F_1 $(1, 58) = 17.89$, $MSE = 120898$, $p < .001$, $\eta_p^2 = .236$	$F_2(1, 22) = 25.95$, $MSE = 52542$, $p < .001$, $\eta_p^2 = .541$
Day x Blocks	F_1 $(4.48, 259.85) = 16.88$, $MSE = 132020$, $p < .001$, $\eta_p^2 = .225$	$F_2(9, 198) = 51.78$, $MSE = 26657$, $p < .001$, $\eta_p^2 = .702$
Group x Blocks	F_1 $(9, 522) = 10.28$, $MSE = 66551$, $p < .001$, $\eta_p^2 = .150$	$F_2(5.64, 124.18) = 45.16$, $MSE = 43014$, $p < .001$, $\eta_p^2 = .672$
Length x Blocks	F_1 $(5.99, 347.46) = 24.32$, $MSE = 43728$, $p < .001$, $\eta_p^2 = .295$	$F_2(9, 198) = 7.07$, $MSE = 11840$, $p < .001$, $\eta_p^2 = .243$
Day x Group x Length	F_1 $(1, 58) = 20.46$, $MSE = 32072$, $p < .001$, $\eta_p^2 = .261$	$F_2(1, 22) = 15.07$, $MSE = 13606$, $p = .001$, $\eta_p^2 = .406$

Day x Group x Blocks	F_1 (9, 522) = 4.86, MSE = 18911, $p < .001$, $\eta_p^2 = .077$	F_2(9, 198) = 17.95, MSE = 7454, $p < .001$, $\eta_p^2 = .449$
Day x Length x Blocks	F_1 (5.94, 344.38) = 4.21, MSE = 7313, $p < .001$, $\eta_p^2 = .068$	F_2(9, 198) = 4.12, MSE = 2120, $p < .001$, $\eta_p^2 = .158$
Group x Length x Blocks	F_1 (9, 522) = 4.31, MSE = 5155, $p < .001$, $\eta_p^2 = .069$	F_2(9, 198) = 3.56, MSE = 2128, $p < .001$, $\eta_p^2 = .139$
Day x Group x Length x Blocks	F_1 (9, 522) = 1.97, MSE = 2258, $p < .05$, $\eta_p^2 = .033$	F_2(9, 198) = 2.34, MSE = 972, $p < .05$, $\eta_p^2 = .096$
Day 1		
Group	F_1 (1, 58) = 30.83, MSE = 6984003, $p < .001$, $\eta_p^2 = .347$	F_2(1, 22) = 1159.82, MSE = 2801264, $p < .001$, $\eta_p^2 = .981$
Length	F_1 (1, 58) = 103.02, MSE = 625368, $p < .001$, $\eta_p^2 = .640$	F_2(1, 22) = 25.98, MSE = 254586, $p < .001$, $\eta_p^2 = .542$
Blocks	F_1 (3.11, 180.35) = 59.81, MSE = 1152122, $p < .001$, $\eta_p^2 = .508$	F_2(5.14, 113.12) = 107.43, MSE = 280758, $p < .001$, $\eta_p^2 = .830$
Group x Length	F_1 (1, 58) = 22.86, MSE = 138753, $p < .001$, $\eta_p^2 = .283$	F_2(1, 22) = 24.76, MSE = 59812, $p < .001$, $\eta_p^2 = .530$
Group x Blocks	F_1 (9, 522) = 11.28, MSE = 75041, $p < .001$, $\eta_p^2 = .163$	F_2(5.19, 114.14) = 42.40, MSE = 52320, $p < .001$, $\eta_p^2 = .658$
Length x Blocks	F_1 (5.41, 313.79) = 19.84, MSE = 46346, $p < .001$, $\eta_p^2 = .255$	F_2(9, 198) = 7.74, MSE = 11555, $p < .001$, $\eta_p^2 = .260$
Group x Length x Blocks	F_1 (9, 522) = 4.07, MSE = 5711, $p < .001$, $\eta_p^2 = .066$	F_2(9, 198) = 3.30, MSE = 2348, $p < .001$, $\eta_p^2 = .130$
Typical adults Day 1		
Length	F_1 (1, 29) = 19.40, MSE = 87490, $p < .001$, $\eta_p^2 = .401$	F_1 (1, 22) = 8.12, MSE = 33800, $p < .01$, $\eta_p^2 = .269$
Blocks	F_1 (2.98, 86.37) = 19.55, MSE = 203021, $p < .001$, $\eta_p^2 = .403$	F_2 (5.37, 118.20) = 53.55, MSE = 45247, $p < .001$, $\eta_p^2 = .709$
Length x Blocks	F_1 (9, 261) = 7.22, MSE = 5276, $p < .001$, $\eta_p^2 = .199$	F_2 (9, 198) = 4.55, MSE = 2294, $p < .001$, $\eta_p^2 = .171$

Bonferroni corrected t-tests(α = .005)		
Block 1	$t_1(29) = 6.38, p < .001$	$t_2(22) = 4.39, p < .001$
Block 2	$t_1(29) = 4.57, p < .001$	$t_2(22) = 2.06, p = .051$
Block 3	$t_1(29) = 5.30, p < .001$	$t_2(22) = 3.88, p < .001$
Block 4	$t_1(29) = 1.39, p = .175$	$t_2(22) = 1.06, p = .301$
Block 5	$t_1(29) = 2.67, p = .012$	$t_2(22) = 1.98, p = .060$
Block 6	$t_1(29) = 0.84, p = .410$	$t_2(22) = 0.88, p = .386$
Block 7	$t_1(29) = 2.05, p = .049$	$t_2(22) = 0.81, p = .428$
Block 8	$t_1(29) = 1.29, p = .209$	$t_2(22) = 0.93, p = .363$
Block 9	$t_1(29) = 1.90, p = .068$	$t_2(22) = 1.40, p = .175$
Block 10	$t_1(29) = 1.28, p = .211$	$t_2(22) = 1.65, p = .113$
Dyslexic Day 1		
Length	F_1 (1, 29) = 88.68, MSE = 676631, $p < .001$, $\eta_p^2 = .754$	F_2 (1, 22) = 34.86, MSE = 280598, $p < .001$, $\eta_p^2 = .613$
Blocks	F_1 (2.94, 85.15) = 41.11, MSE = 1244155, $p < .001$, $\eta_p^2 = .586$	F_2 (5.24, 115.32) = 96.21, MSE = 280814, $p < .001$, $\eta_p^2 = .814$
Length x Blocks	F_1 (4.38, 126.98) = 13.62, MSE = 58157, $p < .001$, $\eta_p^2 = .320$	F_2 (9, 198) = 6.83, MSE = 11609, $p < .001$, $\eta_p^2 = .237$
Bonferroni corrected t-tests(α = .005)		
Block 1	$t_1(29) = 8.47, p < .001$	$t_2(22) = 5.53, p < .001$
Block 2	$t_1(29) = 6.45, p < .001$	$t_2(22) = 4.78, p < .001$
Block 3	$t_1(29) = 5.73, p < .001$	$t_2(22) = 3.32, p = .003$
Block 4	$t_1(29) = 4.81, p < .001$	$t_2(22) = 3.65, p < .001$
Block 5	$t_1(29) = 4.21, p < .001$	$t_2(22) = 3.63, p < .001$
Block 6	$t_1(29) = 2.89, p = .007$	$t_2(22) = 2.59, p = .017$
Block 7	$t_1(29) = 5.34, p < .001$	$t_2(22) = 2.80, p = .010$
Block 8	$t_1(29) = 2.87, p = .008$	$t_2(22) = 1.73, p = .098$
Block 9	$t_1(29) = 4.29, p < .001$	$t_2(22) = 3.74, p < .001$
Block 10	$t_1(29) = 4.29, p < .001$	$t_2(22) = 2.20, p = .038$

Day 7		
Group	F_1 (1, 58) = 15.58, MSE = 2782781, $p < .001$, $\eta_p^2 = .212$	F_2 (1, 22) = 2174.11, MSE = 1115049, $p < .001$, $\eta_p^2 = .990$
Length	F_1 (1, 58) = 34.71, MSE = 78216, $p < .001$, $\eta_p^2 = .374$	F_2 (1, 22) = 6.60, MSE = 32035, $p < .05$, $\eta_p^2 = .231$
Blocks	F_1 (3.33, 193.05) = 21.35, MSE = 214521, $p < .001$, $\eta_p^2 = .269$	F_2 (9, 198) = 45.44, MSE = 31688, $p < .001$, $\eta_p^2 = .674$
Group x Length	F_1 (1, 58) = 6.31, MSE = 14216, $p < .05$, $\eta_p^2 = .098$	F_2 (1, 22) = 12.36, MSE = 6336, $p < .005$, $\eta_p^2 = .360$
Group x Blocks	F_1 (9, 522) = 2.81, MSE = 10421, $p < .005$, $\eta_p^2 = .046$	F_2 (9, 198) = 14.17, MSE = 4271, $p < .001$, $\eta_p^2 = .392$
Length x Blocks	F_1 (9, 522) = 6.46, MSE = 6071, $p < .001$, $\eta_p^2 = .100$	F_2 (9, 198) = 3.45, MSE = 2404, $p < .001$, $\eta_p^2 = .135$
Group x Length x Blocks	F_1 (9, 522) = 1.81, MSE = 1701, $p = .064$, $\eta_p^2 = .030$	F_2 (9, 198) = 2.50, MSE = 752, $p < .01$, $\eta_p^2 = .102$
Control Day 7		
Length	F_1 (1, 29) = 7.52, MSE = 12871, $p < .01$, $\eta_p^2 = .206$	F_2 (1, 22) = 2.04, MSE = 4938, $p = .167$, $\eta_p^2 = .085$
Blocks	F_1 (3.24, 93.85) = 8.42, MSE = 46950, $p < .001$, $\eta_p^2 = .225$	F_2 (9, 198) = 18.14, MSE = 6667, $p < .001$, $\eta_p^2 = .452$
Length x Blocks	F_1 (6.09, 176.72) = 3.25, MSE = 2662, $p < .005$, $\eta_p^2 = .101$	F_2 (9, 198) = 1.92, MSE = 707, $p < .05$, $\eta_p^2 = .080$
Bonferroni corrected t-tests(α = .005)		
Block 1	$t_1(29) = 4.68$, $p < .001$	$t_2(22) = 3.01$, $p = .006$
Block 2	$t_1(29) = 1.33$, $p = .193$	$t_2(22) = 0.66$, $p = .517$
Block 3	$t_1(29) = 0.55$, $p = .589$	$t_2(22) = 0.37$, $p = .716$
Block 4	$t_1(29) = 1.68$, $p = .104$	$t_2(22) = 1.38$, $p = .182$
Block 5	$t_1(29) = 0.28$, $p = .786$	$t_2(22) = 0.09$, $p = .933$
Block 6	$t_1(29) = 0.15$, $p = .883$	$t_2(22) = 0.07$, $p = .945$
Block 7	$t_1(29) = 0.02$, $p = .982$	$t_2(22) = 0.02$, $p = .981$
Block 8	$t_1(29) = 2.00$, $p = .055$	$t_2(22) = 1.96$, $p = .063$
Block 9	$t_1(29) = 1.74$, $p = .093$	$t_2(22) = 1.24$, $p = .227$

Block 10	t_1(29) = 0.71, p = .482	t_2(22) = 0.48, p = .635
Dyslexic Day 7		
Length	F_1 (1, 29) = 28.47, MSE = 79562, p < .001, η_p^2 = .495	F_2 (1, 22) = 11.36, MSE = 33433, p < .005, η_p^2 = .340
Blocks	F_1 (2.84, 82.46) = 13.43, MSE = 230665, p < .001, η_p^2 = .317	F_2 (9, 198) = 46.42, MSE = 29293, p < .001, η_p^2 = .678
Length x Blocks	F_1 (9, 261) = 4.50, MSE = 5970, p < .001, η_p^2 = .134	F_2 (9, 198) = 3.88, MSE = 2449, p < .001, η_p^2 = .150
Bonferroni corrected t-tests(α = .005)		
Block 1	t_1(29) = 5.53, p < .001	t_2(22) = 4.51, p < .001
Block 2	t_1(29) = 3.93, p < .001	t_2 (22) = 4.17, p < .001
Block 3	t_1(29) = 3.02, p < .005	t_2 (22) = 2.22, p = .037
Block 4	t_1(29) = 0.58, p = .568	t_2 (22) = 0.53, p = .605
Block 5	t_1(29) = 1.65, p = .110	t_2 (22) = 1.60, p = .123
Block 6	t_1(29) = 0.46, p = .648	t_2 (22) = 0.61, p = .548
Block 7	t_1(29) = 2.43, p = .021	t_2 (22) = 1.84, p = .079
Block 8	t_1(29) = 1.93, p = .063	t_2 (22) = 1.35, p = .190
Block 9	t_1(29) = 2.28, p = .030	t_2 (22) = 2.05, p = .052
Block 10	t_1(29) = 1.35, p = .188	t_2 (22) = 1.20, p = .245

5 Appendix 5

5.1 Analysis of the result in Experiment 9 of Chapter 6.

5.1.1 Experiment 9

Effects	*By-participants analysis*	*By-items analysis*
NAMING LATENCY (correct, trimmed RTs)		
Overall analysis		
Group	$F_1(1, 48) = 27.27$, $MSE = 11143499$, $p < .001$, $\eta_p^2 = .362$	$F_2(1, 22) = 1805.64$, $MSE = 5409480$, $p < .001$, $\eta_p^2 = .988$
Day	$F_1(1, 48) = 48.83$, $MSE = 2604649$, $p < .001$, $\eta_p^2 = .504$	$F_2(1, 22) = 1475.74$, $MSE = 1297235$, $p < .001$, $\eta_p^2 = .985$
Length	$F_1(1, 48) = 77.71$, $MSE = 542981$, $p < .001$, $\eta_p^2 = .618$	$F_2(1, 22) = 28.09$, $MSE = 273775$, $p < .001$, $\eta_p^2 = .561$
Blocks	$F_1(3.03, 145.37) = 70.02$, $MSE = 912546$, $p < .001$, $\eta_p^2 = .593$	$F_2(5.28, 116.16) = 104.79$, $MSE = 260569$, $p < .001$, $\eta_p^2 = .826$
Day x Group	$F_1(1, 48) = 7.35$, $MSE = 392105$, $p < .01$, $\eta_p^2 = .133$	$F_2(1, 22) = 427.00$, $MSE = 189409$, $p < .001$, $\eta_p^2 = .951$
Day x Length	$F_1(1, 48) = 26.14$, $MSE = 73977$, $p < .001$, $\eta_p^2 = .353$	$F_2(1, 22) = 34.92$, $MSE = 30696$, $p < .001$, $\eta_p^2 = .613$
Group x Length	$F_1(1, 48) = 19.68$, $MSE = 137523$, $p < .001$, $\eta_p^2 = .291$	$F_2(1, 22) = 24.17$, $MSE = 72422$, $p < .001$, $\eta_p^2 =.524$
Day x Blocks	$F_1(5.73, 275.12) = 12.84$, $MSE = 56130$, $p < .001$, $\eta_p^2 = .211$	$F_2(9, 198) = 36.18$, $MSE = 17325$, $p < .001$, $\eta_p^2 = .622$
Group x Blocks	$F_1(9, 432) = 9.97$, $MSE = 43704$, $p < .001$, $\eta_p^2 = .172$	$F_2(4.84, 106.45) = 36.68$, $MSE = 40328$, $p < .001$, $\eta_p^2 = .625$
Length x Blocks	$F_1(6.66, 319.76) = 20.99$, $MSE = 24311$, $p < .001$, $\eta_p^2 = .304$	$F_2(9, 198) = 6.03$, $MSE = 8792$, $p < .001$, $\eta_p^2 = .215$
Day x Group x Length	$F_1(1, 48) = 4.19$, $MSE = 11859$, $p < .05$, $\eta_p^2 = .080$	$F_2(1, 22) = 8.56$, $MSE = 3798$, $p < .01$, $\eta_p^2 = .280$

Day x Group x Blocks	$F_1(9, 432) = 2.54$, *MSE* = 7061, $p < .01$, $\eta_p^2 = .050$	$F_2(5.38, 118.44) = 6.57$, *MSE* = 5272, $p < .001$, $\eta_p^2 = .230$
Day x Length x Blocks	$F_1(6.67, 320.32) = 4.29$, *MSE* = 4621, $p < .001$, $\eta_p^2 = .082$	$F_2(9, 198) = 4.26$, *MSE* = 2039, $p < .001$, $\eta_p^2 = .162$
Group x Length x Blocks	$F_1(9, 432) = 2.57$, *MSE* = 2203, $p < .01$, $\eta_p^2 = .304$	$F_2(9, 198) = 1.69$, *MSE* = 1000, $p = .093$, $\eta_p^2 = .071$
Day x Group x Length x Blocks	$F_1(9, 432) = 0.91$, *MSE* = 725, $p = .518$, $\eta_p^2 = .019$	$F_2(9, 198) = 0.74$, *MSE* = 356, $p = .669$, $\eta_p^2 = .033$
Day 1		
Group	$F_1(1, 48) = 28.58$, *MSE* = 7858119, $p < .001$, $\eta_p^2 = .373$	$F_2(1, 22) = 1610.48$, *MSE* = 3811673, $p < .001$, $\eta_p^2 = .987$
Length	$F_1(1, 48) = 70.19$, *MSE* = 508899, $p < .001$, $\eta_p^2 = .594$	$F_2(1, 22) = 35.38$, *MSE* = 243908, $p < .001$, $\eta_p^2 = .617$
Blocks	$F_1(3.55, 170.20) = 58.21$, *MSE* = 693496, $p < .001$, $\eta_p^2 = .548$	$F_2(5.15, 113.31) = 119.07$, *MSE* = 236136, $p < .001$, $\eta_p^2 = .844$
Group x Length	$F_1(1, 48) = 15.87$, *MSE* = 115075, $p < .001$, $\eta_p^2 = .249$	$F_2(1, 22) = 23.11$, *MSE* = 54695, $p < .001$, $\eta_p^2 = .512$
Group x Blocks	$F_1(9, 432) = 8.46$, *MSE* = 39690, $p < .001$, $\eta_p^2 = .150$	$F_2(4.60, 101.12) = 28.78$, *MSE* = 37601, $p < .001$, $\eta_p^2 = .567$
Length x Blocks	$F_1(6.19, 297.33) = 17.27$, *MSE* = 25047, $p < .001$, $\eta_p^2 = .265$	$F_2(9, 198) = 7.92$, *MSE* = 8984, $p < .001$, $\eta_p^2 = .265$
Group x Length x Blocks	$F_1(9, 432) = 2.30$, *MSE* = 2298, $p < .05$, $\eta_p^2 = .046$	$F_2(9, 198) = 1.62$, *MSE* = 2114, $p = .112$, $\eta_p^2 = .069$
4-letter		
Blocks	$F_1(4.20, 201.51) = 29.26$, *MSE* = 169773, $p < .001$, $\eta_p^2 = .379$	$F_2(3.86, 42.47) = 29.49$, *MSE* = 90163, $p < .001$, $\eta_p^2 = .728$
Group	$F_1(1, 48) = 22.74$, *MSE* = 3035663, $p < .001$, $\eta_p^2 = .321$	$F_2(1, 11) = 500.48$, *MSE* = 1476588, $p < .001$, $\eta_p^2 = .978$
Blocks x Group	$F_1(9, 432) = 4.71$, *MSE* = 12746, $p < .001$, $\eta_p^2 = .089$	$F_2(3.23, 35.53) = 7.72$, *MSE* = 17103, $p < .001$, $\eta_p^2 = .412$

7-letter		
Blocks	$F_1(3.62, 173.87) = 70.77$, $MSE = 524919$, $p < .001$, $\eta_p^2 = .596$	$F_2(4.83, 53.14) = 110.02$, $MSE = 196433$, $p < .001$, $\eta_p^2 = .909$
Group	$F_1(1, 48) = 33.20$, $MSE = 4937532$, $p < .001$, $\eta_p^2 = .409$	$F_2(1, 11) = 1340.15$, $MSE = 2389780$, $p < .001$, $\eta_p^2 = .992$
Blocks x Group	$F_1(9, 432) = 9.79$, $MSE = 29242$, $p < .001$, $\eta_p^2 = .169$	$F_2(9, 99) = 26.25$, $MSE = 23946$, $p < .001$, $\eta_p^2 = .705$
British adults Day 1		
Length	$F_1 (1, 29) = 19.40$, $MSE = 87490$, $p < .001$, $\eta_p^2 = .401$	$F_1 (1, 22) = 8.12$, $MSE = 33800$, $p < .01$, $\eta_p^2 = .269$
Blocks	$F_1 (2.98, 86.37) = 19.55$, $MSE = 203021$, $p < .001$, $\eta_p^2 = .403$	$F_2 (5.37, 118.20) = 53.55$, $MSE = 45247$, $p < .001$, $\eta_p^2 = .709$
Length x Blocks	$F_1 (9, 261) = 7.22$, $MSE = 5276$, $p < .001$, $\eta_p^2 = .199$	$F_2 (9, 198) = 4.55$, $MSE = 2294$, $p < .001$, $\eta_p^2 = .171$
Bonferroni corrected t-tests(α = .005)		
Block 1	$t_1(29) = 6.38$, $p < .001$	$t_2(22) = 4.39$, $p < .001$
Block 2	$t_1(29) = 4.57$, $p < .001$	$t_2(22) = 2.06$, $p = .051$
Block 3	$t_1(29) = 5.30$, $p < .001$	$t_2(22) = 3.88$, $p < .001$
Block 4	$t_1(29) = 1.39$, $p = .175$	$t_2(22) = 1.06$, $p = .301$
Block 5	$t_1(29) = 2.67$, $p = .012$	$t_2(22) = 1.98$, $p = .060$
Block 6	$t_1(29) = 0.84$, $p = .410$	$t_2(22) = 0.88$, $p = .386$
Block 7	$t_1(29) = 2.05$, $p = .049$	$t_2(22) = 0.81$, $p = .428$
Block 8	$t_1(29) = 1.29$, $p = .209$	$t_2(22) = 0.93$, $p = .363$
Block 9	$t_1(29) = 1.90$, $p = .068$	$t_2(22) = 1.40$, $p = .175$
Block 10	$t_1(29) = 1.28$, $p = .211$	$t_2(22) = 1.65$, $p = .113$
Block 1 and 10 comparison (α = 0.025)		
4-letter	$t_1(29) = 4.68$, $p < .001$	$t_2(11) = 9.20$, $p < .001$
7-letter	$t_1(29) = 6.14$, $p < .001$	$t_2(11) = 13.20$, $p < .001$

Chinese adult Day 1		
Length	$F_1(1, 19) = 40.39$, $MSE = 461652$, $p < .001$, $\eta_p^2 = .680$	F_2 $(1, 22) = 51.95$, $MSE = 264803$, $p < .001$, $\eta_p^2 = .703$
Blocks	F_1 $(3.19, 60.59) = 32.66$, $MSE = 609574$, $p < .001$, $\eta_p^2 = .632$	F_2 $(4.60, 101.28) = 98.12$, $MSE = 248920$, $p < .001$, $\eta_p^2 = .817$
Length x Blocks	F_1 $(4.35, 82.62) = 9.07$, $MSE = 26415$, $p < .001$, $\eta_p^2 = .323$	F_2 $(9, 198) = 5.99$, $MSE = 7770$, $p < .001$, $\eta_p^2 = .214$
Bonferroni corrected t-tests($\alpha = .005$)		
Block 1	$t_1(19) = 8.51$, $p < .001$	$t_2(22) = 5.24$, $p < .001$
Block 2	$t_1(19) = 5.33$, $p < .001$	$t_2(22) = 5.81$, $p < .001$
Block 3	$t_1(19) = 5.08$, $p < .001$	$t_2(22) = 4.65$, $p < .001$
Block 4	$t_1(19) = 2.54$, $p = .020$	$t_2(22) = 2.45$, $p = .023$
Block 5	$t_1(19) = 3.75$, $p < .001$	$t_2(22) = 3.34$, $p = .003$
Block 6	$t_1(19) = 2.98$, $p = .008$	$t_2(22) = 3.56$, $p = .002$
Block 7	$t_1(19) = 3.16$, $p = .005$	$t_2(22) = 2.63$, $p = .015$
Block 8	$t_1(19) = 4.23$, $p < .001$	$t_2(22) = 3.74$, $p < .001$
Block 9	$t_1(19) = 4.18$, $p < .001$	$t_2(22) = 5.57$, $p < .001$
Block 10	$t_1(19) = 3.93$, $p < .001$	$t_2(22) = 3.98$, $p < .001$
Block 1 and 10 comparison ($\alpha = 0.025$)		
4-letter	$t_1(19) = 7.38$, $p < .001$	$t_2(11) = 7.71$, $p < .001$
7-letter	$t_1(19) = 8.57$, $p < .001$	$t_2(11) = 14.94$, $p < .001$
Day 7		
Group	F_1 $(1, 48) = 19.67$, $MSE = 3677485$, $p < .001$, $\eta_p^2 = .291$	F_2 $(1, 22) = 1666.16$, $MSE = 1787217$, $p < .001$, $\eta_p^2 = .987$
Length	F_1 $(1, 48) = 42.09$, $MSE = 108059$, $p < .001$, $\eta_p^2 = .467$	$F_2(1, 22) = 16.23$, $MSE = 60563$, $p < .001$, $\eta_p^2 = .425$
Blocks	F_1 $(4.49, 215.42) = 28.12$, $MSE = 139559$, $p < .001$, $\eta_p^2 = .369$	$F_2(9, 198) = 43.67$, $MSE = 35052$, $p < .001$, $\eta_p^2 = .665$
Group x Length	F_1 $(1, 48) = 13.36$, $MSE = 34307$, $p < .001$, $\eta_p^2 = .218$	$F_2(1, 22) = 20.07$, $MSE = 21525$, $p < .001$, $\eta_p^2 = .477$

Group x Blocks	F_1 (9, 432) = 4.47, MSE = 11075, $p < .001$, $\eta_p^2 = .085$	F_2(9, 198) = 13.95, MSE = 5631, $p < .001$, $\eta_p^2 = .388$
Length x Blocks	F_1 (6.89, 330.94) = 6.36, MSE = 5460, $p < .001$, $\eta_p^2 = .117$	F_2(9, 198) = 2.30, MSE = 1847, $p < .05$, $\eta_p^2 = .095$
Group x Length x Blocks	F_1 (9, 432) = 0.96, MSE = 630, $p = .474$, $\eta_p^2 = .020$	F_2(9, 198) = 0.69, MSE = 277, $p = .721$, $\eta_p^2 = .030$
4-letter		
Blocks	F_1 (5.17, 248.23) = 15.91, MSE = 39669, $p < .001$, $\eta_p^2 = .249$	F_2 (9, 99) = 18.75, MSE = 12022, $p < .001$, $\eta_p^2 = .630$
Group	F_1 (1, 48) = 16.07, MSE = 1500702, $p < .001$, $\eta_p^2 = .251$	**F_2** (1, 11) = 724.24, MSE = 708235, $p < .001$, $\eta_p^2 = .985$
Blocks x Group	F_1 (9, 432) = 3.07, MSE = 4394, $p < .001$, $\eta_p^2 = .060$	F_2 (9, 99) = 5.51, MSE = 2343, $p < .001$, $\eta_p^2 = .334$
7-letter		
Blocks	F_1 (4.95, 237.56) = 30.00, MSE = 92710, $p < .001$, $\eta_p^2 = .385$	F_2 (9, 99) = 25.80, MSE = 24876, $p < .001$, $\eta_p^2 = .701$
Group	F_1 (1, 48) = 23.00, MSE = 2211090, $p < .001$, $\eta_p^2 = .324$	F_2 (1, 11) = 942.68, MSE = 1100507, $p < .001$, $\eta_p^2 = .988$
Blocks x Group	F_1 (9, 432) = 4.30, MSE = 7311, $p < .001$, $\eta_p^2 = .082$	F_2 (9, 99) = 9.34, MSE = 3565, $p < .001$, $\eta_p^2 = .459$
British adults Day 7		
Length	F_1 (1, 29) = 7.52, MSE = 12871, $p < .01$, $\eta_p^2 = .206$	F_2 (1, 22) = 2.04, MSE = 4938, $p = .167$, $\eta_p^2 = .085$
Blocks	F_1 (3.24, 93.85) = 8.42, MSE = 46950, $p < .001$, $\eta_p^2 = .225$	F_2 (9, 198) = 18.14, MSE = 6667, $p < .001$, $\eta_p^2 = .452$
Length x Blocks	F_1 (6.09, 176.72) = 3.25, MSE = 2662, $p < .005$, $\eta_p^2 = .101$	F_2 (9, 198) = 1.92, MSE = 707, $p < .05$, $\eta_p^2 = .080$

Bonferroni corrected t-tests(α = .005)		
Block 1	$t_1(29) = 4.68, p < .001$	$t_2(22) = 3.01, p = .006$
Block 2	$t_1(29) = 1.33, p = .193$	$t_2(22) = 0.66, p = .517$
Block 3	$t_1(29) = 0.55, p = .589$	$t_2(22) = 0.37, p = .716$
Block 4	$t_1(29) = 1.68, p = .104$	$t_2(22) = 1.38, p = .182$
Block 5	$t_1(29) = 0.28, p = .786$	$t_2(22) = 0.09, p = .933$
Block 6	$t_1(29) = 0.15, p = .883$	$t_2(22) = 0.07, p = .945$
Block 7	$t_1(29) = 0.02, p = .982$	$t_2(22) = 0.02, p = .981$
Block 8	$t_1(29) = 2.00, p = .055$	$t_2(22) = 1.96, p = .063$
Block 9	$t_1(29) = 1.74, p = .093$	$t_2(22) = 1.24, p = .227$
Block 10	$t_1(29) = 0.71, p = .482$	$t_2(22) = 0.48, p = .635$
Block 1 and 10 comparison (α = 0.025)		
4-letter	$t_1(29) = 3.66, p < .001$	$t_2(11) = 5.63, p < .001$
7-letter	$t_1(29) = 5.26, p < .001$	$t_2(11) = 8.89, p < .001$
Chinese adults Day 7		
Length	F_1 (1, 19) = 28.42, MSE = 110058, $p < .001$, $\eta_p^2 = .599$	F_2 (1, 22) = 32.36, MSE = 77149, $p < .001$, $\eta_p^2 = .595$
Blocks	F_1 (3.87, 73.43) = 17.53, MSE = 130337, $p < .001$, $\eta_p^2 = .480$	F_2 (9, 198) = 40.56, MSE = 34016, $p < .001$, $\eta_p^2 = .648$
Length x Blocks	F_1 (4.59, 87.23) = 3.45, MSE = 5506, $p < .01$, $\eta_p^2 = .154$	F_2 (9, 198) = 1.69, MSE = 1416, $p = .094$, $\eta_p^2 = .071$
Bonferroni corrected t-tests(α = .005)		
Block 1	$t_1(19) = 5.71, p < .001$	$t_2(22) = 3.70, p < .001$
Block 2	$t_1(19) = 3.71, p = .002$	$t_2(22) = 4.23, p < .001$
Block 3	$t_1(19) = 4.49, p < .001$	$t_2(22) = 2.73, p = .012$
Block 4	$t_1(19) = 3.81, p < .001$	$t_2(22) = 3.64, p < .001$
Block 5	$t_1(19) = 4.08, p < .001$	$t_2(22) = 2.11, p = .046$
Block 6	$t_1(19) = 0.86, p = .400$	$t_2(22) = 0.87, p = .395$
Block 7	$t_1(19) = 2.74, p = .013$	$t_2(22) = 2.40, p = .054$
Block 8	$t_1(19) = 2.84, p = .010$	$t_2(22) = 3.14, p = .005$
Block 9	$t_1(19) = 1.86, p = .079$	$t_2(22) = 3.47, p = .002$
Block 10	$t_1(19) = 2.63, p = .016$	$t_2(22) = 2.85, p = .009$

Block 1 and 10 comparison (α = 0.025)		
4-letter	$t_1(19) = 5.46$, $p < .001$	$t_2(11) = 7.02$, $p < .001$
7-letter	$t_1(19) = 8.12$, $p < .001$	$t_2(11) = 9.58$, $p < .001$

6 Appendix: Stimulus of each experiment

6.1 All the stimulus in each set

Item	Length	Set A	Set B	Set C	Set D	Set E	Set F	Set G	Set H	Set I	Set J	Set K	Set L
barg	1	*	*				*						
blop	1			*				*					*
brin	1				*				*				
brup	1					*				*			
carg	1					*							*
cark	1		*										
clat	1	*		*				*					
cont	1												*
cran	1				*				*				
cugg	1									*			
dast	1												*
delp	1									*			
dift	1	*		*				*					
drap	1		*				*						
dreb	1					*							
drof	1				*				*				
goom	1	*		*				*					
gort	1									*			
grol	1		*				*						
gulb	1				*				*				
jant	1		*				*						
jeph	1					*							
jesh	1				*				*				
jice	1	*		*				*					
julk	1									*			
kelf	1				*				*				
kess	1									*			
kest	1												*
kilp	1	*		*				*					
krin	1		*				*						
larn	1		*				*						
leng	1	*		*				*					
loke	1				*				*				
lont	1					*				*			
marb	1												*
munt	1					*							

nart	1									*			
nate	1					*							*
nipe	1	*	*				*						
nirl	1			*				*					
nool	1				*				*				
pite	1												*
plid	1												*
plin	1					*							
quap	1				*				*				
quen	1									*			
quib	1		*				*						
quoz	1	*		*				*					
reen	1									*			
relb	1					*							
reld	1				*				*				
rell	1												*
rint	1	*	*				*						
roke	1			*				*					
tife	1									*			
tond	1												*
torp	1				*				*				
trok	1					*							
trom	1	*		*				*					
turb	1		*				*						
varb	1					*							
wost	1												*
yerp	1									*			
ying	1	*											
yint	1		*				*						
yost	1				*				*				
yulf	1			*				*					
zort	1					*							
bencort	2												*
blispod	2		*			*	*				*		
bloffey	2	*		*				*				*	
brellet	2				*				*				
bruckep	2									*			
cantoom	2		*				*				*		
carklin	2												*
carmunt	2			*				*			*		
coftrip	2				*	*			*			*	*

cushlap	2	*								*		*	
dempton	2												*
despale	2	*	*				*					*	
dintorn	2				*				*				
doldorm	2			*				*			*		
drentcy	2					*							
drestel	2									*			
gamroon	2		*				*				*		
gatlern	2				*				*				
gromple	2			*				*					
gurmint	2	*								*		*	
jaggert	2	*	*				*					*	
jerslaw	2									*			
jespord	2			*				*			*		
jimplen	2				*				*				
joshule	2					*							
keffert	2	*	*				*					*	
kintore	2												*
krallep	2			*				*					
krendle	2									*			
kusherm	2				*				*		*		
lagrole	2									*		*	
larquof	2					*							
lentwin	2				*				*		*		
lintone	2	*	*				*					*	
lonnart	2			*				*			*		
marpoon	2												*
mattoch	2					*							
nasheet	2									*	*		
nelpoon	2					*							*
nessale	2			*				*					
nolfern	2	*	*				*					*	
nultorp	2				*				*				
pembert	2												*
plinore	2												*
pronnet	2					*							
quarple	2				*				*				
querpid	2		*				*				*		
questal	2	*								*		*	
quovent	2			*				*					
rashtin	2	*			*				*			*	

roffler	2					*							*
rotgroy	2									*		*	
runstle	2			*				*			*		
ruskeng	2		*				*				*		
tarbish	2	*		*				*				*	
teadert	2									*			
tismole	2		*				*						*
trimsol	2					*							
trockle	2				*				*		*		
vushood	2					*							
wedrick	2												*
yapsote	2				*				*				
yarchin	2		*				*						
yebbler	2	*								*		*	
yorquin	2			*				*			*		
zadroon	2					*							

6.2 The distribution of sets in each experiment

Chapter	Experiment	Set(s)
2	1	A
	2	B, C, D
	3	E
3	4	F, G, H, I
	5	G, H, I
	6	J, K
4	7	L
5	8	E
6	9	E

6.3 The stimuli for the orthographic distractors in Experiment 6

	orthographic distractors	Fragments from real words
1	guab	arguably
2	eopl	people
3	ivit	creativity
4	avel	travel
5	nsig	insight
6	reci	appreciation
7	dapt	adaption
8	iori	priorities
9	ustr	frustrated
10	ctic	practical
11	rsto	understood
12	efin	refine
13	luti	solution
14	iump	triumph
15	inst	brainstorm
16	imit	limit
17	uery	query
18	amou	famous
19	ater	material
20	trai	constraint
21	inat	imagination

22	plor	exploration
23	ctua	intellectual
24	rchi	architecture
25	elie	believe
26	mmed	immediate
27	ogra	program
28	rict	restrict
29	iver	diversity
30	rcep	perception
31	oduc	produce
32	eaut	beauty
33	ient	orientation
34	armo	harmony
35	lega	elegance
36	litu	solitude
37	plif	uplift
38	bser	observe
39	aliz	realize
40	uali	quality
41	asur	measure
42	ysic	physics
43	ythm	rhythm
44	escr	description
45	ngua	language

46	erta	entertainment
47	lleg	college
48	onas	monastic
49	etai	retailer
50	stor	history
51	muni	community
52	ghte	frighten
53	uris	tourist
54	mome	moment
55	ucce	success
56	tten	attention
57	ndiv	individual
58	oces	process
59	uran	insurance
60	peri	experience
61	amil	family
62	atte	matter
63	bjec	objection
64	uara	guarantee
65	ojec	projector
66	urro	surround
67	xtre	extreme
68	rmon	harmony
69	mati	information

70	rhap	perhaps
71	iter	literature
72	hoos	choose
73	erwi	otherwise
74	anda	standard
75	icul	difficult
76	scov	discover
77	ltim	ultimate
78	alua	valuable
79	ccep	accept
80	esul	result
81	tuit	intuition
82	cisi	decision
83	geme	management
84	pref	preference

6.4 Words stimulus used in Experiment 8

High frequency words		Low frequency words	
4-letter	7-letter	4-letter	7-letter
deal	darling	deed	default
king	kitchen	kite	ketchup
news	nervous	nest	neutral
team	teacher	tart	toaster
beat	believe	bake	biscuit
card	country	cord	concert
mark	machine	mute	mermaid
pick	promise	pier	profile
rest	respect	ripe	rubbish
Wear	welcome	wolf	wealthy
cost	contact	cart	concise
poor	perhaps	plug	perfume

Reference

Adams, M. J. (1990). *Beginning to read: Learning and thinking about print*. Cambridge. MA: MIT Press.

Alloway, T. (2007). The Automatic Working Memory Assessment (AWMA)[Computer Software]. London: Harcourt Assessment.

Andrews, S. (1982). Phonological recoding: Is the regularity effect consistent? *Memory & Cognition, 10*, 565-575.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 802-814.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 234-254.

Archer, N., & Bryant, P. (2001). Investigating the role of context in learning to read: A direct test of Goodman's model. *British Journal of Psychology, 92*(4), 579-591.

Atkins, P. W., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*(4), 537-552.

Avons, S., Wragg, C. A., Cupples, L., & Lovegrove, W. J. (1998). Measures of phonological short-term memory and their relationship to vocabulary development. *Applied Psycholinguistics, 19,* 583-602.

Baayen, R. H. (2009). LanguageR: datasets and functions with 'Analyzing Linguistic' Data: A practical introduction to statistics'. R package version 0.9555.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database [CD-ROM]: Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Baayen, R. H., Piepenbrock, R., & Van Rijn, H. (1995). The CELEX lexical database, release 2 [CD-ROM]. : Philadelphia: Linguistic Data Consortium & University of Pennsylvania.

Baddeley, A. (1986). *Working memory*. Oxford: Clarendon Press.

Baddeley, A. (1998). Recent developments in working memory. *Current Opinion in Neurobiology, 8*(2), 234-238.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A., Della Sala, S., Gray, C., Papagno, C., & Spinnler, H. (1997). Testing central executive functioning with a Pencil-and-paper test. In P. Rabbitt (Ed.), *Methodology of frontal and executive functions* (pp. 61-80). Hove, England: Psychology Press.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158-173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The Psychology of Learning and Motivation, 8,* 47-89.

Baker, C., & Jones, S. P. (1998). *Encyclopedia of bilingualism and bilingual education*. Clevedon, United Kingdom: Multilingual Matters.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*(2), 283-316.

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: beyond measures of central tendency. *Journal of Experimental Psychology: General, 128*(1), 32-55.

Bates, D. M. (2007). *Linear Mixed Model Implementation in Lme4*. Unpublished manuscript, Department of Statistics, University of Wisconsin-Madison, Madison, WI.

Bates, D. M., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. Retrieved from http: //cran.R-project.org/package=lme4(R package version 0.999999-0).

Ben-Dror, I., Pollatsek, A., & Scarpati, S. (1991). Word identification in isolation and in context by college dyslexic students. *Brain and Language, 40*(4), 471-490.

Bialystok, E., McBride-Chang, C., & Luk, G. (2005). Bilingualism, Language Proficiency, and Learning to Read in Two Writing Systems. *Journal of Educational Psychology, 97*(4), 580-590.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers, 33*(1), 73-79.

Blazely, A. M., Coltheart, M., & Casey, B. J. (2005). Semantic impairment with and without surface dyslexia: Implications for models of reading. *Cognitive Neuropsychology, 22*(6), 695-717.

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Interfering neighbours: the impact of novel word learning on the identification of visually similar words. *Cognition, 97*(3), 45-54.

Bowers, P. G., Steffy, R., & Tate, E. (1988). Comparison of the effects of IQ control methods on memory and naming speed predictors of reading disability. *Reading Research Quarterly, 23,* 304-309.

Bowey, J. A., & Hansen, J. (1994). The development of orthographic rimes as units of word recognition. *Journal of Experimental Child Psychology, 58*(3), 465-488.

Bowey, J. A., & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading: A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 92*(3), 203-219.

Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., Götz, K., & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry, 49*(4), 422-432.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301,* 419-421.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary reading skill differences in young adults. *Journal of Learning Disabilities, 40*(3), 226-243.

Breitenstein, C., Kamping, S., Jansen, A., Schomacher, M., & Knecht, S. (2004). Word learning can be achieved without feedback: implications for aphasia therapy. *Restorative Neurology and Neuroscience, 22*(6), 445-458.

Breitenstein, C., & Knecht, S. (2002). Development and validation of a language learning model for behavioral and functional-imaging studies. *Journal of Neuroscience Methods, 114*(2), 173-179.

Brinton, L. J. (2002). Grammaticalization versus lexicalization reconsidered. In T. Fanego & M. G. López-Cous, Javier P (Eds.), *English Historical Syntax and Morphology* (pp. 67-97). Amsterdam: John Benjamins.

Brooks, L. (1977). Non-analytic correspondences and pattern in word pronunciation. In J. Renquin (Ed.), *Attention and Performance* (pp. 163-177). Hillsdale, New Jersey: Erlbaum.

Brown, G. D., & Hulme, C. (1996). Nonword repetition, STM, and word age-of-acquisition: A computational model. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 129-148). Hove: LEA.

Bruck, M. (1990). Word-recognition skills of adults with childhood diagnoses of dyslexia. *Developmental Psychology, 26*(3), 439-454.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology, 28*(5), 874-886.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Experimental psychology, 58*(5), 412-424.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977-990.

Bürki, A., Spinelli, E., & Gaskell, G. M. (2012). A written word is worth a thousand spoken words: The influence of spelling on spoken-word production. *Journal of Memory and Language, 67*(4), 449-467.

Byrne, B., & Fielding-Barnsley, R. (1989). Phonemic awareness and letter knowledge in the child's acquisition of the alphabetic principle. *Journal of Educational Psychology, 81*(3), 313-321.

Callens, M., Tops, W., & Brysbaert, M. (2012). Cognitive profile of students who enter higher education with an indication of dyslexia. *PloS one, 7*(6), e38081.

Castles, A., Holmes, V., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56*, 445-468.

Castles, A., & Nation, K. (2008). Learning to be a good orthographic reader. *Journal of Research in Reading, 31*(1), 1-7.

Chalmers, K. A., & Burt, J. S. (2008). Phonological and semantic information in adults' orthographic learning. *Acta Psychologica, 128*(1), 162-175.

Chen, Y., Fu, S., Iversen, S. D., Smith, S. M., & Matthews, P. M. (2002). Testing for dual brain processing routes in reading: a direct contrast of Chinese character and pinyin reading using fMRI. *Journal of Cognitive Neuroscience, 14*(7), 1088-1098.

Cheung, H. (1995). Predicting pronunciation accuracy in a second language: The role of phonological awareness. *Psychologia: An International Journal of Psychology in the Orient, 38*(4), 266-274.

Cheung, H. (1999). Improving phonological awareness and word reading in a later learned alphabetic script. *Cognition, 70*(1), 1-26.

Cheung, H., Chen, H. C., Lai, C. Y., Wong, O. C., & Hills, M. (2001). The development of phonological awareness: Effects of spoken language experience and orthography. *Cognition, 81*(3), 227-241.

Clay, F., Bowers, J. S., Davis, C. J., & Hanley, D. A. (2007). Teaching adults new words: the role of practice and consolidation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(5), 970-976.

Coady, J. (1979). A psycholinguistic model of the ESL reader. In R. Mackay, B. Barkman & R. R. Jordan (Eds.), *Reading in a second language* (pp. 5-12). Rowley, MA: Newbury House.

Coane, J. H., & Balota, D. A. (2010). Repetition priming across distinct contexts: Effects of lexical status, word frequency, and retrieval test. *The Quarterly Journal of Experimental Psychology, 63*(12), 2376-2398.

Collier, V. P. (1989). How long? A synthesis of research on academic achievement in a second language. *TESOL quarterly, 23*(3), 509-531.

Colombo, L., Pasini, M., & Balota, D. A. (2006). Dissociating the influence of familiarity and meaningfulness from word frequency in naming and lexical decision performance. *Memory & Cognition, 34*(6), 1312-1324.

Coltheart, M. (2005). Analysing developmental disorders of reading. *Australian Journal of Learning Difficulties, 10*(3-4), 53-62.

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human perception and performance, 20*(6), 1197-1211.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256.

Cook, V. J., & Bassetti, B. (2005). *Second language writing systems*. Clevedon, United Kingdom: Multilingual Matters.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: an analysis of 2,342 words. *The Quarterly Journal of Experimental Psychology, 60*(8), 1072-1082.

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *The Quarterly Journal of Experimental Psychology, 66*(5), 946-972.

Cosky, M. J. (1976). The role of letter recognition in word recognition. *Memory & Cognition, 4*(2), 207-214.

Crystal, D. (1997). *English as a global language*. Cambridge, United Kingdom: Cambridge University Press.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of educational Research, 49*(2), 222-251.

Cunningham, A. E. (2006). Accounting for children's orthographic learning while reading text: Do children self-teach? *Journal of Experimental Child Psychology, 95*(1), 56-77.

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: examining the role of self-teaching. *Journal of Experimental Child Psychology, 82*(3), 185-199.

Davies, R., Rodríguez-Ferreiro, J., Suárez, P., & Cuetos, F. (2013). Lexical and sub-lexical effects on accuracy, reaction time and response duration: impaired and typical word and pseudoword reading in a transparent orthography. *Reading and Writing, 26*(5), 721-738.

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1536), 3773-3800.

De Jong, P. F., Bitter, D. J., van Setten, M., & Marinus, E. (2009). Does phonological recoding occur during silent reading, and is it necessary for orthographic learning? *Journal of Experimental Child Psychology, 104*(3), 267-282.

de Jong, P. F., & Messbauer, V. (2011). Orthographic context and the acquisition of orthographic knowledge in normal and dyslexic readers. *Dyslexia, 17*(2), 107-122.

De Luca, M., Barca, L., Burani, C., & Zoccolotti, P. (2008). The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Cognitive and Behavioral Neurology, 21*(4), 227-235.

Deacon, S. H., Cook, K., & Parrila, R. (2012). Identifying high-functioning dyslexics: is self-report of early reading problems enough? *Annals of dyslexia*, *62*(2), 120-134.

Delaney, P. F., Reder, L. M., Staszewski, J. J., & Ritter, F. E. (1998). The strategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science, 9*(1), 1-7.

Della Sala, S., Gray, C., Baddeley, A., Allamano, N., & Wilson, L. (1999). Pattern span: a tool for unwelding visuo–spatial memory. *Neuropsychologia, 37*(10), 1189-1199.

Dodd, B., Campbell, R., & Worrall, L. (1996). *Evaluating theories of language: evidence from disordered communication*. London: Whurr Publishers.

Duchek, J. M., & Neely, J. H. (1989). A dissociative word-frequency X levels-of-processing interaction in episodic recognition and lexical decision tasks. *Memory & Cognition, 17*(2), 148-162.

Duff, F. J., & Hulme, C. (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading, 16*(6), 504-525.

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35-39.

Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of reading and arithmetic skills in 7-to 10-year-olds. *Journal of Experimental Child Psychology, 91*(2), 113-136.

Duyck, W., Desmet, T., Verbeke, L. P., & Brysbaert, M. (2004). WordGen: a tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers, 36*(3), 488-499.

Ehri, L. C. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities, 22*(6), 356-365.

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. Gough, L. C. Ehri & R. Treiman (Eds.), *Reading acquisition* (pp. 107-143). Hillsdale, NJ: Lawrence Erlbaum.

Ehri, L. C. (1998). Research on learning to read and spell: A personal-historical perspective. *Scientific Studies of Reading, 2*(2), 97-114.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167-188.

Ehri, L. C., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing, 7*(3), 295-326.

Ehri, L. C., & Wilce, L. S. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology, 71*(1), 26-40.

Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology, 46*(4), 375-384.

Ellis, A. W., Ferreira, R., Cathles-Hagan, P., Holt, K., Jarvis, L., & Barca, L. (2009). Word learning and the cerebral hemispheres: from serial to parallel processing of written words. *Philosophical Transactions of the Royal Society Biological Sciences, 364*(1536), 3675-3696.

Fan, K., Gao, J., & Ao, X. (1984). Pronunciation principles of the Chinese character and alphabetic writing scripts. *Chinese character reform, 3*, 23-27.

Ferrand, L. (2000). Reading aloud polysyllabic words and nonwords: The syllabic length effect reexamined. *Psychonomic Bulletin & Review, 7*(1), 142-148.

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983a). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General, 112*(3), 309-347.

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983b). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General, 112*(3), 309-346.

Field, A. (2009). *Discovering statistics using SPSS*. London: Sage.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. London: Sage Publications Limted.

Firth, U. (1972). *Components of reading disability* (Unpublished doctoral dissertation), University of New South Wales, Australia.

Forbach, G. B., Stanners, R. F., & Hochhaus, L. (1974). Repetition and practice effects in a lexical decision task. *Memory & Cognition, 2*(2), 337-339.

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*(6), 627-635.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 680-698.

Forster, K. I., Mohan, K., Hector, J., Kinoshita, S., & Lupker, S. (2003). The mechanics of masked priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state of the art* (pp. 3-37). Philadelphia: Psychology Press.

Frith, U. (1986). A developmental framework for developmental dyslexia. *Annals of dyslexia, 36*(1), 67-81.

Gallagher, A., Laxon, V., Armstrong, E., & Frith, U. (1996). Phonological difficulties in high-functioning dyslexics. *Reading and Writing, 8*(6), 499-509.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition*, *89*(2), 105-132.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513-543.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200-213.

Gathercole, S. E., Hitch, G. J., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33*(6), 966-979.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*(1), 65-77.

Geva, E., & Wang, M. (2001). The development of basic reading skills in children: A cross-language perspective. *Annual Review of Applied Linguistics*, *21*, 182-204.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1166-1183.

Gombert, J. E. (1992). *Metalinguistic development*. Chicago: University of Chicago Press.

Grainger, J., & Jacobs, A. M. (1996). Orthographic processing in visual word recognition: a multiple read-out model. *Psychological Review, 103*(3), 518-565.

Griffiths, Y. M., & Snowling, M. J. (2002). Predictors of exception word and nonword reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology, 94*(1), 34-43.

Grossberg, S., & Stone, G. (1986). Neural dynamics of word recognition and recall: attentional priming, learning, and resonance. *Psychological Review, 93*(1), 46-74.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology: Section A*, *56*(7), 1213-1236.

Hachmann, W. M., Bogaerts, L., Szmalec, A., Woumans, E., Duyck, W., & Job, R. (2014). Short-term memory for order but not for item information is impaired in developmental dyslexia. *Annals of dyslexia*, *47*, 1-16.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: insights from connectionist models. *Psychological Review, 106*(3), 491-528.

Hatcher, J., Snowling, M. J., & Griffiths, Y. M. (2002). Cognitive assessment of dyslexic students in higher education. *British Journal of Educational Psychology, 72*(1), 119-133.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65*(1), 41-57.

Hauptmann, B., & Karni, A. (2002). From primed to learn: the saturation of repetition priming and the induction of long-term memory. *Cognitive Brain Research, 13*(3), 313-322.

Havelka, J., & Frankish, C. (2010). Is RoAsT tougher than StEAk?: The effect of case mixing on perception of multi-letter graphemes. *Psihologija, 43*(1), 103-116.

Henderson, L., Weighall, A., Brown, H., & Gaskell, G. (2013). Online lexical competition during spoken word recognition and word learning in children and adults. *Child Development, 84*(5), 1668-1685.

Herrmann, J. A., Matyas, T., & Pratt, C. (2006). Meta-analysis of the nonword reading deficit in Specific Reading Disorder. *Dyslexia, 12*(3), 195-221.

Hershenson, M., & Haber, R. N. (1965). Role of meaning in the perception of briefly exposed words. *Canadian Journal of Psychology/Revue canadienne de psychologie, 19*(1), 42-46.

Hino, Y., & Lupker, S. J. (2000). Effects of word frequency and spelling–to–sound regularity in naming with and without preceding lexical decision. *Journal of Experimental Psychology: Human perception and performance, 26*(1), 166-183.

Ho, C. S. H., & Bryant, P. (1997). Phonological skills are important in learning to read Chinese. *Developmental Psychology, 33*(6), 946-951.

Ho, C. S. H., Ng, T. T., & Ng, W. K. (2003). A “radical” approach to reading development in Chinese: The role of semantic radicals and phonetic radicals. *Journal of Literacy Research, 35*(3), 849-878.

Hogaboam, T. W., & Perfetti, C. A. (1978). Reading skill and the role of verbal experience in decoding. *Journal of Educational Psychology, 70*(5), 717-729.

Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition, 59*(2), 119-147.

Holmes, V. M., & Babauta, M. L. (2005). Single or dual representations for reading and spelling? *Reading and Writing, 18*(3), 257-280.

Hu, C. F. (2008). Use orthography in L2 auditory word learning: Who benefits? *Reading and Writing, 21*(8), 823-841.

Huang, H. S., & Hanley, J. R. (1995). Phonological awareness and visual skills in learning to read Chinese and English. *Cognition, 54*(1), 73-98.

Hulme, C., Goetz, K., Gooch, D., Adams, J., & Snowling, M. J. (2007). Paired-associate learning, phoneme awareness, and learning to read. *Journal of Experimental Child Psychology, 96*(2), 150-166.

Hulme, C., & Snowling, M. J. (2014). The interface between spoken and written language: developmental disorders. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1634), 1-8.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5*(1), 67-104.

Jared, D., Mcrae, K., & Seidenberg, M. S. (1990). The basis of consistency effects in word naming. *Journal of Memory and Language, 29*(6), 687-715.

Jarrold, C., Thorn, A. S., & Stephens, E. (2009). The relationships among verbal short-term memory, phonological awareness, and new word learning: Evidence from typical development and Down syndrome. *Journal of Experimental Child Psychology, 102*(2), 196-218.

Johnston, M., McKague, M., & Pratt, C. (2004). Evidence for an automatic orthographic code in the processing of visually novel word forms. *Language and Cognitive Processes, 19*(2), 273-317.

Jones, M. W., Branigan, H. P., & Kelly, M. L. (2008). Visual deficits in developmental dyslexia: relationships between non-linguistic visual tasks and their contribution to components of reading. *Dyslexia, 14*(2), 95-115.

Juphard, A., Carbonnel, S., & Valdois, S. (2004). Length effect in reading and lexical decision: evidence from skilled readers and a developmental dyslexic participant. *Brain and Cognition, 55*(2), 332-340.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. New York, United States: Cambridge University Press.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978-990.

Kyte, C. S., & Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology, 93*(2), 166-185.

Lam, A. S., Perfetti, C. A., & Bell, L. (1991). Automatic phonetic transfer in bidialectal reading. *Applied Psycholinguistics, 12*(03), 299-311.

Landi, N., Perfetti, C. A., Bolger, D. J., Dunlap, S., & Foorman, B. R. (2006). The role of discourse context in developing word form representations: A paradoxical relation between reading and learning. *Journal of Experimental Child Psychology, 94*(2), 114-133.

Lavidor, M., & Ellis, A. W. (2002). Word length and orthographic neighborhood size effects in the left and right cerebral hemispheres. *Brain and Language, 80*(1), 45-62.

Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: when adults learn new words. *Cognitive Psychology, 55*(4), 306-353.

Lee, C. H. (1999). A locus of the word-length effect on word recognition. *Reading Psychology, 20*(2), 129-150.

Lee, T., Lo, W. K., Ching, P., & Meng, H. (2002). Spoken language resources for Cantonese speech processing. *Speech Communication, 36*(3), 327-342.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(01), 1-38.

Levy, B. A., Abello, B., & Lysynchuk, L. (1997). Transfer from word training to reading in context: Gains in reading fluency and comprehension. *Learning Disability Quarterly, 20*(3), 173-188.

Li, T., McBride-Chang, C., Wong, A., & Shu, H. (2012). Longitudinal predictors of spelling and reading comprehension in Chinese as an L1 and English as an L2 in Hong Kong Chinese children. *Journal of Educational Psychology, 104*(2), 286-301.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22*(1), 1-35.

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 883-914.

Lovett, M. C., Daily, L. Z., & Reder, L. M. (2000). A source activation theory of working memory: Cross-task prediction of performance in ACT-R. *Cognitive Systems Research, 1*(2), 99-118.

Lupker, S. J., Brown, P., & Colombo, L. (1997). Strategic control in a naming task: Changing routes or changing deadlines? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(3), 570-590.

Lupker, S. J., Kinoshita, S., Coltheart, M., & Taylor, T. E. (2003). Mixing costs and mixing benefits in naming words, pictures, and sums. *Journal of Memory and Language*, *49*(4), 556-575.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: studies with artificial lexicons. *Journal of Experimental Psychology: General*, *132*(2), 202-227.

Majerus, S., & Boukebza, C. (2013). Short-term memory for serial order supports vocabulary development: New evidence from a novel word learning paradigm. *Journal of Experimental Child Psychology, 116*(4), 811-828.

Majerus, S., Poncelet, M., Van der Linden, M., Albouy, G., Salmon, E., Sterpenich, V., Vandewalle, G., Collette, F., & Maquet, P. (2006). The left intraparietal sulcus and verbal short-term memory: Focus of attention or serial order? *Neuroimage, 32*(2), 880-891.

Maloney, E., Risko, E. F., O'Malley, S., & Besner, D. (2009). Tracking the transition from sublexical to lexical processing: on the creation of orthographic and phonological lexical representations. *Quarterly Journal of Experimental Psychology, 62*(5), 858-867.

Manis, F. R. (1985). Acquisition of word identification skills in normal and disabled readers. *Journal of Educational Psychology, 77*(1), 78-90.

Marinus, E., & de Jong, P. F. (2010). Variability in the word-reading performance of dyslexic readers: Effects of letter length, phoneme length and digraph presence. *Cortex*, *46*(10), 1259-1271.

Martelli, M., De Luca, M., Lami, L., Pizzoli, C., Pontillo, M., Spinelli, D., & Zoccolotti, P. (2014). Bridging the gap between different measures of the reading speed deficit in developmental dyslexia. *Experimental brain research, 232*(1), 237-252.

Martin-Chang, S. L., Levy, B. A., & O’Neil, S. (2007). Word acquisition, retention, and transfer: Findings from contextual and isolated word training. *Journal of Experimental Child Psychology, 96*(1), 37-56.

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*(03), 379-413.

Mason, M. (1978). From print to sound in mature readers as a function of reader ability and two forms of orthographic regularity. *Memory & Cognition*, *6*(5), 568-581.

Mayor, J., & Plunkett, K. (2010). A neurocomputational account of taxonomic responding and fast mapping in early word learning. *Psychological Review, 117*(1), 1-31.

Mayringer, H., & Wimmer, H. (2000). Pseudoname learning by German-speaking children with dyslexia: Evidence for a phonological learning deficit. *Journal of Experimental Child Psychology, 75*(2), 116-133.

McBride-Chang, C., Cheung, H., Chow, B. W. Y., Chow, C. S. L., & Choi, L. (2006). Metalinguistic skills and vocabulary knowledge in Chinese (L1) and English (L2). *Reading and Writing, 19*(7), 695-716.

McBride-Chang, C., & Ho, C. S. H. (2005). Predictors of beginning reading in Chinese and English: A 2-year longitudinal study of Chinese kindergartners. *Scientific Studies of Reading, 9*(2), 117-144.

McBride-Chang, C., & Manis, F. R. (1996). Structural invariance in the associations of naming speed, phonological awareness, and verbal reasoning in good and poor readers: A test of the double deficit hypothesis. *Reading and Writing, 8*(4), 323-339.

McBride-Chang, C., & Treiman, R. (2003). Hong Kong Chinese kindergartners learn to read English analytically. *Psychological Science*, *14*(2), 138-143.

McBride-Chang, C., Wagner, R. K., Muse, A., Chow, B. W. Y., & Shu, H. U. A. (2005). The role of morphological awareness in children's vocabulary acquisition in English. *Applied Psycholinguistics*, *26*(3), 415-435.

McCann, R. S., & Besner, D. (1987). Reading pseudohomophones: Implications for models of pronunciation assembly and the locus of word-frequency effects in naming. *Journal of Experimental Psychology: Human perception and performance, 13*(1), 14-24.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419-457.

McKague, M., Davis, C., Pratt, C., & Johnston, M. B. (2008). The role of feedback from phonology to orthography in orthographic learning: an extension of item-based accounts. *Journal of Research in Reading, 31*(1), 55-76.

McKague, M., Pratt, C., & Johnston, M. B. (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition, 80*(3), 231-262.

McKay, A., Davis, C., Savage, G., & Castles, A. (2008). Semantic involvement in reading aloud: evidence from a nonword training study. *Journal of Experimental Psychology: Learning Memory and Cognition, 34*(6), 1495-1517.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review, 119*(4), 831-877.

Melby-Lervåg, M., Lervåg, A., Lyster, S.-A. H., Klem, M., Hagtvet, B., & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children’s vocabulary development. *Psychological Science, 23*(10), 1092-1098.

Menghini, D., Finzi, A., Carlesimo, G. A., & Vicari, S. (2011). Working memory impairment in children with developmental dyslexia: is it just a phonological deficity? *Developmental Neuropsychology, 36*(2), 199-213.

Merkx, M., Rastle, K., & Davis, M. H. (2011). The acquisition of morphological knowledge investigated through artificial language learning. *The Quarterly Journal of Experimental Psychology, 64*(6), 1200-1220.

Michas, I. C., & Henry, L. A. (1994). The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology, 12*(2), 147-163.

Moll, K., Hulme, C., Nag, S., & Snowling, M. J. (2013). Sentence repetition as a marker of language skills in children with dyslexia. *Applied Psycholinguistics, 34*, 1-19.

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling--sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(1), 183-206.

Monaghan, P., & Ellis, A. W. (2010). Modeling reading development: Cumulative, incremental learning in a computational model of word naming. *Journal of Memory and Language, 63*(4), 506-525.

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology, 91*, 167-180.

Mortimore, T., & Crozier, W. R. (2006). Dyslexia and difficulties with study skills in higher education. *Studies in Higher Education, 31*(2), 235-251.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Developmental Psychology, 40*(5), 665-681.

Myers, R. H. (1990). *Classical and modern regression with applications*. Boston, Mass: PWS and Kent Publishing Company.

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. C. McKeown, M. (Ed.), *The nature of vocabulary acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Nation, K. (2008). Learning to read words. *The Quarterly Journal of Experimental Psychology, 61*(8), 1121-1133.

Nation, K., & Cocksey, J. (2009). The relationship between knowing a word and reading it aloud in children's word reading development. *Journal of Experimental Child Psychology, 103*(3), 296-308.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*(4), 342-356.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage, and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition, and Pedagogy* (pp. 4-19). Cambridge: Cambridge University Press.

Nelson, J. R., Balass, M., & Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Language & Literacy, 8*(2), 25-44.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review, 13*(1), 45-52.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: a new French lexical database. *Behavior Research Methods, Instruments, & Computers, 36*(3), 516-524.

Newell, A., & Rosenbloom, P. S. (1981). *Mechanisms of skill acquisition and the law of practice* (Vol. 1). Hillsdale, New Jersey: Lawrence Erlbaum Associates

Nosarti, C., Mechelli, A., Green, D. W., & Price, C. J. (2009). The impact of second language learning on semantic and nonsemantic first language reading. *Cerebral Cortex*, bhp101.

O'Brien, B. A., Van Orden, G. C., & Pennington, B. F. (2013). Do dyslexics misread a ROWS for a ROSE? *Reading and Writing*, *26*(3), 381-402.

O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2011). Complementary learning systems. *Cognitive Science, 35*, 1-20.

Olague, H. M., Etzkorn, L. H., Gholston, S., & Quattlebaum, S. (2007). Empirical validation of three software metrics suites to predict fault-proneness of object-oriented classes developed using highly iterative or agile software development processes. *Software Engineering, IEEE Transactions on, 33*(6), 402-419.

Oldfield, R. C. (1966). Things, words and the brain. *The Quarterly Journal of Experimental Psychology, 18*(4), 340-353.

Ouellette, G. P. (2006). What's meaning got to do with it: the role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554-566.

Pagliuca, G., Arduino, L. S., Barca, L., & Burani, C. (2008). Fully transparent orthography, yet lexical reading aloud: The lexicality effect in Italian. *Language and Cognitive Processes*, *23*(3), 422-433.

Paizi, D., De Luca, M., Zoccolotti, P., & Burani, C. (2013). A comprehensive evaluation of lexical reading in Italian developmental dyslexics. *Journal of Research in Reading, 3*, 303-329.

Paizi, D., Zoccolotti, P., & Burani, C. (2011). Lexical stress assignment in Italian developmental dyslexia. *Reading and Writing, 24*(4), 443-461.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *The Quarterly Journal of Experimental Psychology, 48*(1), 98-107.

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes, 22*(5), 700-726.

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line orthographic influences on spoken language in a semantic task. *Journal of Cognitive Neuroscience, 21*(1), 169-179.

Peereman, R., Content, A., & Bonin, P. (1998). Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language, 39*(2), 151-174.

Pennington, B. F. (1990). The genetics of dyslexia. *Journal of Child Psychology and Psychiatry, 31*(2), 193-201.

Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development, 72*(3), 816-833.

Pennington, B. F., Lefly, D. L., Van Orden, G. C., Bookman, M. O., & Smith, S. D. (1987). Is phonology bypassed in normal or dyslexic development? *Annals of dyslexia*, *37*(1), 62-89.

Perea, M., & Carreiras, M. (1996). Efectos de frecuencia silábica y vecindad ortográfica en la pronunciación de palabras y pseudopalabras. *Psicológica, 17,* 425-440.

Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human perception and performance, 24*(1), 134-144.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri & R. Treiman (Eds.), *Reading acquisition*. Hillsdale, New Jersey: Erlbaum.

Perfetti, C. A. (1997). Sentences, individual differences, and multiple texts: Three issues in text comprehension. *Discourse Processes, 23*(3), 337-355.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language, 30*(4), 473-485.

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27*(1), 59-70.

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. Gorfien (Ed.), *On the consequences of meaning selection* (pp. 67-86). Washington, DC: American Psychological Association.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In V. Ludo, E. Carsten & R. Pieter (Eds.), *Precursors of Functional Literacy* (pp. 67-86). Amsterdam, Philadelphia: John Benjamins.

Perfetti, C. A., & Tan, L. H. (1999). The constituency model of Chinese word identification. In A. Inhoff, H. S. Chen & J. Wang (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 115-134). Mahwah, New Jersey: Lawrence Erlbaum.

Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 24-33.

Perry, C., & Ziegler, J. C. (2002). Cross-language computational investigation of the length effect in reading aloud. *Journal of Experimental Psychology: Human perception and performance, 28*(4), 990-1001.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review, 114*(2), 273-315.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology, 61*(2), 106-151.

Petrova, A., Gaskell, M. G., & Ferrand, L. (2011). Orthographic consistency and word-frequency effects in auditory word recognition: new evidence from lexical decision and rime detection. *Frontiers in Psychology, 2,* 1-11.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review, 103*(1), 56-115.

Raaijmakers, J. G., Schrijnemakers, J., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": Common misconceptions and alternative solutions. *Journal of Memory and Language, 41*(3), 416-426.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly, 27*(1), 29-53.

Rastle, K., & Coltheart, M. (1998). Whammies and double whammies: The effect of length on nonword reading. *Psychonomic Bulletin & Review, 5*(2), 277-282.

Rastle, K., Kinoshita, S., Lupker, S. J., & Coltheart, M. (2003). Cross-task strategic effects. *Memory & Cognition, 31*(6), 867-876.

Rastle, K., McCormick, S. F., Bayliss, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1588-1594.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81*(2), 275-280.

Reid, A. A., Szczerbinski, M., Iskierka-Kasperek, E., & Hansen, P. (2007). Cognitive profiles of adult developmental dyslexics: Theoretical implications. *Dyslexia, 13*(1), 1-24.

Reitsma, P. (1983). Printed word learning in beginning readers. *Journal of Experimental Child Psychology, 36*(2), 321-339.

Reynolds, M., Mulatti, C., & Besner, D. (2012). Reading Nonwords Aloud: Evidence for dynamic control in skilled readers. *Psychonomic Bulletin & Review, 19*(6), 1135-1141.

Richards, L. G., & Heller, F. (1976). Recognition thresholds as a function of word length. *The American Journal of Psychology, 89*, 455-466.

Ricketts, J., Bishop, D. V., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *The Quarterly Journal of Experimental Psychology, 62*(10), 1948-1966.

Rose, J. (2009). Identifying and teaching children and young people with dyslexia and literacy difficulties. Retrieved from: www.teachernet.gov.uk/wholeschool/sen

Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology, 100*(1), 175-191.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature, 323*(6088), 533-536.

Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review, 81*(2), 99-118.

Salasoo, A., Shiffrin, R. M., & Feustel, T. C. (1985). Building permanent memory codes - codification and repetition effects in word identification. *Journal of Experimental Psychology: General, 114*(1), 50-77.

Sandak, R., Mencl, W. E., Frost, S. J., & Pugh, K. R. (2004). The neurobiological basis of skilled and impaired reading: Recent findings and new directions. *Scientific Studies of Reading, 8*(3), 273-292.

Savage, R. S., Frederickson, N., Goodwin, R., Patni, U., Smith, N., & Tuersley, L. (2005). Relationships among rapid digit naming, phonological processing, motor automaticity, and speech perception in poor, average, and good readers and spellers. *Journal of Learning Disabilities, 38*(1), 12-28.

Savin, H. B. (1972). What the child knows about speech when he starts to learn to read. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relatinship between speech and reading*. Cambridge, MA: MIT Press.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human perception and performance, 3*(1), 1-17.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. Shapiro, P. Accardo & A. Capute (Eds.), *Specific Reading Disability: A View of the Spectrum* (pp. 75-119). Timonium, MD: York Press.

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*(6), 1270-1281.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime 1.0. . Pittsburgh, PA: Psychological Software Tools.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523-568.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition. *Journal of Verbal Learning and Verbal Behavior, 23*(3), 383-404.

Semrud-Clikeman, M., Guy, K., Griffin, J. D., & Hynd, G. W. (2000). Rapid naming deficits in children and adolescents with reading disabilities and attention deficit hyperactivity disorder. *Brain and Language, 74*(1), 70-83.

Ševa, N., Monaghan, P., & Arciuli, J. (2009). Stressing what is important: Orthographic cues and lexical stress assignment. *Journal of Neurolinguistics, 22*(3), 237-249.

Share, D. L. (1995). Phonological recoding and self-teaching - sine-qua-non of reading acquisition. *Cognition, 55*(2), 151-218.

Share, D. L. (1999). Phonological recoding and orthographic learning: a direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*(2), 95-129.

Share, D. L. (2004). Orthographic learning at a glance: on the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology, 87*(4), 267-298.

Share, D. L., & Shalev, C. (2004). Self-teaching in normal and disabled readers. *Reading and Writing, 17*(7-8), 769-800.

Shek, D. T., & Ma, C. (2011). Impact of the Project PATHS in the junior secondary school years: individual growth curve analyses. *The Scientific World Journal, 11*, 253-266.

Sio, U. N., Monaghan, P., & Ormerod, T. (2013). Sleep on it, but only if it is difficult: Effects of sleep on problem solving. *Memory & Cognition, 41*(2), 159-166.

Siok, W. T., & Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Developmental Psychology, 37*(6), 886-899.

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human perception and performance, 24*(1), 185-203.

Smith-Spark, J., Fisk, J., Fawcett, A., & Nicolson, R. (2003). Investigating the central executive in adult dyslexics: Evidence from phonological and visuospatial working memory performance. *European Journal of Cognitive Psychology, 15*(4), 567-587.

Smith-Spark, J. H., & Fisk, J. E. (2007). Working memory functioning in developmental dyslexia. *Memory, 15*(1), 34-56.

Snowling, M., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 346*(1315), 21-27.

Snowling, M., Nation, K., Moxham, P., Gallagher, A., & Frith, U. (1997). Phonological processing skills of dyslexic students in higher education: A preliminary report. *Journal of Research in Reading, 20*(1), 31-41.

Snowling, M. J. (2000). *Dyslexia*. Oxford: Blackwell Publishing.

Snowling, M. J. (2001). From language to reading and dyslexia1. *Dyslexia, 7*(1), 37-46.

Snowling, M. J., & Hulme, C. (2008). *The science of reading: A handbook*. London: John Wiley & Sons.

Solomon, R. L., & Postman, L. (1952). Frequency of usage as a determinant of recognition thresholds for words. *Journal of Experimental Psychology, 43*(3), 195-201.

Stanovich, K. E. (1993). Romance and reality. *Reading Teacher, 47*, 280-280.

Stark, C. E., & McClelland, J. L. (2000). Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(4), 945-972.

Stone, G. O., Vanhoy, M., & Orden, G. C. V. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. *Journal of Memory and Language, 36*(3), 337-359.

Storkel, H. L., Armbruster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research, 49*(6), 1175-1192.

Swain, M. (1981). Bilingual education for majority and minority language children. *Studia linguistica, 35*, 15-32.

Swanson, H. L., & Hsieh, C.-J. (2009). Reading disabilities in adults: A selective meta-analysis of the literature. *Review of educational Research, 79*(4), 1362-1390.

Swanson, H. L., Zheng, X., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities a selective meta-analysis of the literature. *Journal of Learning Disabilities, 42*(3), 260-287.

Taft, M., & Chen, H. C. (1992). Judging homophony in Chinese: The influence of tones. *Advances in psychology, 90*, 151-172.

Taft, M., & Zhu, X. (1995). The representation of bound morphemes in the lexicon: A Chinese study *Morphological aspects of language processing* (pp. 293-316). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Tamminen, J., & Gaskell, M. G. (2008). Newly learned spoken words show long-term lexical competition effects. *The Quarterly Journal of Experimental Psychology, 61*(3), 361-371.

Tamminen, J., Payne, J. D., Stickgold, R., Wamsley, E. J., & Gaskell, M. G. (2010). Sleep spindle activity is associated with the integration of new memories and existing knowledge. *The Journal of Neuroscience, 30*(43), 14356-14360.

Tarone, E. E. (1990). On variation interlanguage: a response to Gregg. *Applied Linguistics, 11*(4), 392-400.

Taylor, J. S. H., Plunkett, K., & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: an artificial orthography paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition, 37*(1), 60-76.

Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: a time-criterion account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(1), 117-138.

Thomson, J. M., & Goswami, U. (2010). Learning novel phonological representations in developmental dyslexia: associations with basic auditory processing of rise time and phonological awareness. *Reading and Writing, 23*(5), 453-473.

Tong, X., Tong, X., & McBride-Chang, C. (2013). A Tale of Two Writing Systems: double dissociation and metalinguistic transfer between chinese and english word reading among Hong Kong children. *Journal of Learning Disabilities, 46*, 1-16.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*(4), 579-593.

Treisman, A. M. (1961). *Attention and speech.* (Unpublished doctoral dissertation), University of Oxford, Oxford, England.

Unsworth, S. J., & Pexman, P. M. (2003). The impact of reader skill on phonological processing in visual word recognition. *The Quarterly Journal of Experimental Psychology: Section A, 56*(1), 63-81.

Valdois, S., Carbonnel, S., Juphard, A., Baciu, M., Ans, B., Peyrin, C., & Segebarth, C. (2006). Polysyllabic pseudo-word processing in reading and lexical decision: converging evidence from behavioral data, connectionist simulations and functional MRI. *Brain Research, 1085*(1), 149-162.

Van den Broeck, W., & Geudens, A. (2012). Old and new ways to study characteristics of reading disability: The case of the nonword-reading deficit. *Cognitive Psychology, 65*(3), 414-456.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*(1), 2-40.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11*(1), 3-32.

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes, 19*(1), 57-95.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*(2), 192-212.

Wagner, R. K., Torgesen, J. K., Laughon, P., Simmons, K., & Rashotte, C. A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology, 85*(1), 83-103.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: Pro-ED.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. *Developmental Psychology, 33*(3), 468-479.

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition, 87*(2), 129-149.

Wang, M., Park, Y., & Lee, K. R. (2006). Korean-English biliteracy acquisition: Cross-language phonological and orthographic transfer. *Journal of Educational Psychology, 98*(1), 148-158.

Wang, M., Perfetti, C. A., & Liu, Y. (2005). Chinese–English biliteracy acquisition: Cross-language and writing system transfer. *Cognition, 97*(1), 67-88.

Warmington, M., Hitch, G. J., & Gathercole, S. E. (2013a). Improving word learning in children using an errorless technique. *Journal of Experimental Child Psychology, 114*(3), 456-465.

Warmington, M., Stothard, S. E., & Snowling, M. J. (2013b). Assessing dyslexia in higher education: the York adult assessment battery-revised. *Journal of Research in Special Educational Needs, 13*(1), 48-56.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. New York: The Psychological Corporation.

Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 50*(2), 439-456.

Wiggs, C. L., Weisberg, J., & Martin, A. (2006). Repetition priming across the adult lifespan—the long and short of it. *Aging, Neuropsychology, and Cognition, 13*(3-4), 308-325.

Wilkinson, G. S., & Robertson, G. (2006). Wide Range Achievement Test (WRAT) 4. Lutz: Psychological Assessment Resources.

Wimmer, H. (1996). The nonword reading deficit in developmental dyslexia: Evidence from children learning to read German. *Journal of Experimental Child Psychology, 61*(1), 80-90.

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., & Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *Journal of Speech, Language, and Hearing Research, 50*(4), 1093-1109.

Wixted, J. T., & Carpenter, S. K. (2007). The Wickelgren power law and the Ebbinghaus savings function. *Psychological Science, 18*(2), 133-134.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science, 2*(6), 409-415.

Wolf, M. (1997). A provisional, integrative account of phonological and naming-speed deficits in dyslexia: Implications for diagnosis and intervention. In B. Blachman (Ed.)*, Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 67-92). Hilsdale, NY: Erlbaum.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*(3), 415-438.

Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-speed processes, timing, and reading A conceptual review. *Journal of Learning Disabilities, 33*(4), 387-407.

Wolff, U. (2009). Phonological and surface subtypes among university students with dyslexia. *International Journal of Disability, Development and Education, 56*(1), 73-91.

Wood, F. B., & Felton, R. H. (1994). Separate linguistic and attentional factors in the development of reading. *Topics in Language Disorders, 14*(4), 42-57.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*(4), 502-529.

Yeung, S. S., & Chan, C. K. (2013). Phonological awareness and oral language proficiency in learning to read English among Chinese kindergarten children in Hong Kong. *British Journal of Educational Psychology, 83*(4), 550-568.

Yeung, S. S., Siegel, L. S., & Chan, C. K. (2013). Effects of a phonological awareness program on English reading and spelling among Hong Kong Chinese ESL children. *Reading and Writing, 26*(5), 681-704.

Young, A., & Bowers, P. G. (1995). Individual difference and text difficulty determinants of reading fluency and expressiveness. *Journal of Experimental Child Psychology, 60*(3), 428-454.

Young, A. R., Beitchman, J. H., Johnson, C., Douglas, L., Atkinson, L., Escobar, M., & Wilson, B. (2002). Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. *Journal of Child Psychology and Psychiatry, 43*(5), 635-645.

Zeegers, P. (2004). Student learning in higher education: A path analysis of academic achievement in science. *Higher Education Research & Development, 23*(1), 35-56.

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review, 5*(4), 683-689.

Ziegler, J. C., Montant, M., & Jacobs, A. M. (1997). The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language, 37*(4), 533-554.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12*(5), 379-384.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Korne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86*(3), 169-193.

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: Where do we stand after more than a decade? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(3), 643-661.

Zoccolotti, P., De Luca, M., Di Filippo, G., Judica, A., & Martelli, M. (2009). Reading development in an orthographically regular language: Effects of length, frequency, lexicality and global processing ability. *Reading and Writing, 22*(9), 1053-1079.

Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and Language, 93*(3), 369-373.

Zorzi, M. (2010). The connectionist dual process (CDP) approach to modelling reading aloud. *European Journal of Cognitive Psychology, 22*(5), 836-860.