The Pennsylvania State University

The Graduate School

College of the Liberal Arts

LEXICAL PARSING STRATEGIES IN TWO LANGUAGES:

CONSTRAINTS ON LANGUAGE SELECTION IN WORD RECOGNITION

A Thesis in

Psychology

by

Bianca M. Sumutka

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2003

The thesis of Bianca Sumutka has been reviewed and approved* by the following:

Judith F. Kroll
Professor of Psychology and Applied Linguistics
Thesis Adviser
Chair of Committee

Paola E. Dussias
Assistant Professor of Spanish and Applied Linguistics

Rick O. Gilmore
Assistant Professor of Psychology

David A. Rosenbaum
Distinguished Professor of Psychology

Kevin R. Murphy
Professor of Psychology
Head of the Department of Psychology

*Signatures are on file in the Graduate School

ABSTRACT

Cross-linguistic research reveals distinct lexical parsing strategies in speakers whose languages differ in spelling-to-sound consistency. These findings tend to support models like the Dual Route Cascaded Model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) in which lexical access occurs via two routes, one in which phonology is directly retrieved in parallel, and a second in which phonology is computed serially. This model predicts that consistent languages rely more on the computational route, whereas less consistent languages rely more on the direct route. If lexical parsing preferences require that languages differing in consistency engage distinct processing, for bilinguals, two languages that differ in consistency should function independently. However, bilingual word recognition research demonstrates that both languages are active and competing, even when they differ in consistency, a finding captured by the Bilingual Interactive Activation model (Dijkstra & van Heuven, 1998). This creates a paradox for word recognition in that cross-language interactions occur, but for lexical parsing, both languages appear to function independently. To examine this paradox, three experiments were performed using tasks shown to reveal parsing preferences. In each experiment, the performance of native English speakers and proficient non-native speakers of English was compared in English. In Experiment 1, participants made lexical decisions to stimuli including pseudohomophones, to compare native and non-native spelling-to-sound computation. In Experiment 2, participants performed clustered lexical decision in which word and nonword stimuli were parsed after the onset or the first vowel to compare the importance of the rime unit in processing. Finally in Experiment 3 participants named words and nonwords including English-German cognates to examine how second

language processing strategies differ for German-English bilinguals in the presence of strong first language cues.

In Experiments 1 and 2, results were similar for native English and native German speakers. In Experiment 1, native Japanese speakers also produced similar results in English, although their performance was more dependent on English proficiency. However when native German speakers named English words in Experiment 3, both first and second language processing strategies were revealed. Implications for current models of word recognition are discussed, and a bilingual Dual Route model is proposed.

TABLE OF CONTENTS

LIST OF FIGURES .... **viii**

LIST OF TABLES .... **xii**

ACKNOWLEDGEMENTS .... **xv**

CHAPTER 1: BACKGROUND AND INTRODUCTION **1**

Cross-linguistic differences in language processing 3
Evidence on syntax 3
Cross-linguistic studies of parsing. 3
The Competition Model 4
Evidence on lexical transfer 8
Constraints on parsing: evidence from phonology 11
Evidence for unique cross-linguistic lexical processing 13
The Dual Route Model. 15
Extending the Dual Route Model to German 17
Cross-language interactions in bilingual word recognition 19
The BIA Model. 22
The BIA+ Model. 24
The Paradox 26

CHAPTER 2: METHODS AND MODELS **28**

A Tale of Two Models 29
The Dual-Route Cascaded Model (DRC). 29
The Bilingual Interactive Activation + (BIA+) Model 37
General Method and Design 37
Participants 38
Experiment 1: Lexical Decision with Pseudohomophones 39
Experiment 2: Cluster lexical decision 41
Experiment 3: Cognate Naming 42
Individual Difference Measures 44
Summary of Proposed Experiments and Tasks 46
Methods of Analysis and General Predictions for Experiments 1-3 45

CHAPTER 3: EXPERIMENT 1: LEXICAL DECISION WITH PSEUDOHOMOPHONES **50**

Method 54
Participants 54
Materials 55
Procedure 55
Results and discussion 56
Participant exclusion and language history questionnaire data 57
Reading span results 59
Calculation of outliers 61

Native English Speaker Group and Comparison to Goswami et al. (2001) Child Data............ 61
Reaction Time for nonwords............ 61
Accuracy for nonwords............ 63
Reaction time for words............ 65
Accuracy for words............ 66
Comparison to Goswami et al. (2001)............ 68
Native German Speaker Group............ 72
Reaction Time data for nonwords............ 72
Accuracy for nonwords............ 74
Reaction time for words............ 77
Accuracy for words............ 78
Native Japanese Speaker Group............ 79
Reaction Time for nonword data............ 79
Accuracy for nonwords............ 83
Reaction time for word data............ 85
Accuracy for words............ 87
Analysis across all language groups............ 89
Analysis across language groups: Syllable length............ 89
Reaction time for nonwords.............. 89
Nonword Accuracy.............. 91
Word Reaction Time.............. 93
Word Accuracy.............. 96
Analysis across language groups: Nonword type............ 98
Reaction Time.............. 98
Accuracy.............. 101
Native German analysis by memory span............ 103
GENERAL DISCUSSION............ 104

CHAPTER 4: EXAMINING PARSING THROUGH CLUSTERED LEXICAL DECISION............ 106

DESCRIPTION OF THE TASK............ 107
METHOD............ 109
Participants............ 109
Materials............ 109
PROCEDURE............ 110
RESULTS AND DISCUSSION............ 111
Exclusion of Participants............ 111
Native English speaker group............ 112
Reaction time for words by phoneme and parse............ 112
Reaction time for nonwords by phoneme and parse............ 117
Reaction Time comparison by SOA for words............ 122
Reaction Time comparison by SOA for nonwords............ 125
Accuracy for words by phoneme and parse............ 128
Accuracy by phoneme and parse for nonwords............ 132
Accuracy comparison by SOA for words............ 136
Accuracy comparison by SOA for nonwords............ 139

Summary of Native English results 141
Native German Speaker Group 142
Exclusion of items and replacement of missing condition means 142
Reaction time for words by phoneme and parse 143
Reaction time comparison by SOA for words 146
Reaction time comparison by method for nonwords 149
Reaction time comparison by SOA for nonwords 150
Accuracy comparison by method for words 151
Accuracy comparison by SOA for words 152
Accuracy comparison by method for nonwords 154
Accuracy comparison by SOA for nonwords 155
Summary of native German results 156
Direct comparison of native English and German groups by method 157
Direct comparison of native English and German groups by SOA 160
GENERAL DISCUSSION 161

CHAPTER 5: WORD LENGTH AND BODY NEIGHBOR EFFECTS IN NAMING 164

METHOD 166
Participants 166
Materials 166
Procedure 167
Results and Discussion 167
Exclusion of participants 168
Calculation of outliers 168
Native English and Native German body x length comparison: word RT 169
Native English and Native German body x length comparison: nonword RT 175
Native English and Native German body x length comparison: word accuracy 180
Native English and Native German body x length comparison: nonword accuracy 184
GENERAL DISCUSSION 187

CHAPTER 6: GENERAL DISCUSSION 192

SUMMARY OF FINDINGS 193
LEXICAL PARSING AND CURRENT MODELS OF WORD RECOGNITION 199
CONCLUSIONS 204

REFERENCES 209

APPENDIX A: *WORD AND NONWORD STIMULI FOR EXPERIMENT 1* **..... 215**

APPENDIX B: *WORD AND NONWORD STIMULI FOR EXPERIMENT 2* **..... 216**

APPENDIX C: *WORD AND NONWORD STIMULI FOR EXPERIMENT 3* **..... 217**

LIST OF FIGURES

Figure 1.1. The Revised Hierarchical Model (adapted from Kroll & Stewart, 1994)…….9

Figure 1.2. The DRC Model (adapted from Coltheart et al., 2001).……………………...16

Figure 1.3. The BIA model of bilingual word recognition (adapted from Dijkstra & van Heuven, 1998)…………………………………………………………..22

Figure 1.4. The BIA+ model of bilingual word recognition (adapted from Dijkstra & van Heuven, 2002)…………………………………………………………..24

Figure 2.1. A bilingual DRC model with two independent processing systems for L1 and L2………………………………………………………………………….31

Figure 2.2. A bilingual DRC Model with L2 representations in the direct route………..32

Figure 2.3. A bilingual DRC Model with L2 representations in the direct and assembled routes……………………………………………………………………….33

Figure 2.4. Cluster lexical decision task for Experiment 2………………………………42

Figure 3.1. Mean reaction times (ms) for Native English speaker group by number of syllables and nonword type for Experiment 1………………………………..63

Figure 3.2. Mean accuracy for Native English speaker group by syllable length and nonword type for Experiment 1………………………………………………65

Figure 3.3. Mean RT data for nonword type collapsed over syllables for adults in Experiment 1 and native English speaking children (Goswami et al., 2001)...70

Figure 3.4. Mean Accuracy collapsed over syllables for adults in Experiment 1 and native English speaking children by nonword type (Goswami et al., 2001)………...71

Figure 3.5. Mean RT for Native German speakers by syllable length and nonword type.……………………………………………………………………….74

Figure 3.6. Mean Accuracy for Native German speakers by syllable length and nonword type.……………………………………………………………………….75

Figure 3.7. Mean RT for high proficiency Native Japanese speakers by syllable length and nonword type………………………………………………………..82

Figure 3.8. Mean RT for low proficiency Native Japanese speakers by syllable length and nonword type………………………………………………………..82

Figure 3.9. Mean Accuracy for Native Japanese speakers by syllable length and nonword type…………………………………………………………………………86

Figure 3.10. Mean RT (ms) by syllable length and language group for nonwords……...90

Figure 3.11. Mean accuracy by syllable length and language group for nonwords……..92

Figure 3.12. Mean RT (ms) by syllable length and language group for words………….95

Figure 3.13. Mean accuracy by syllable length and language group for words…………97

Figure 3.14. Mean RT by nonword type and language group………………………….100

Figure 3.15. Mean RT by nonword type and language group………………………….102

Figure 4.1. Events within the cluster lexical decision task in Experiment 2…………...108

Figure 4.2. Reaction time (ms) for type of parse by number of phonemes for the space method of presentation………………………………………………………113

Figure 4.3. Reaction time (ms) for type of parse by number of phonemes for the asterisks method of presentation……………………………………………………115

Figure 4.4. Reaction time (ms) for words by type of parse and number of phonemes for Treiman and Chafetz (1987)……………………………………………...116

Figure 4.5. Reaction time (ms) for type of parse and number of phonemes for nonwords for the space method of presentation……………………………………..119

Figure 4.6. Reaction time (ms) for type of parse and number of phonemes for nonwords for the asterisks method of presentation…………………………………120

Figure 4.7. Reaction time (ms) for nonwords by type of parse and number of phonemes for Treiman and Chafetz (1987)………………………………………….121

Figure 4.8. Accuracy for type of parse and number of phonemes for words for the space method of presentation……………………………………………………129

Figure 4.9. Accuracy for type of parse and number of phonemes for words for the asterisks method of presentation……………………………………………130

Figure 4.10. Accuracy for words by type of parse and number of phonemes for Treiman and Chafetz (1987)………………………………………………………..131

Figure 4.11. Accuracy for nonwords by type of parse and number of phonemes for the space method of presentation……………………………………………..133

Figure 4.12. Accuracy for nonwords by type of parse and number of phonemes for the asterisks method of presentation…………………………………………..135

Figure 4.13. Accuracy for nonwords by type of parse and number of phonemes for Treiman and Chafetz (1987)……………………………………………136

Figure 4.14. Word reaction time (ms) for native German speakers by type of parse and number of phonemes for the space method of presentation……………….144

Figure 4.15. Word reaction time (ms) for native German speakers by type of parse and number of phonemes for the space method of presentation……………….145

Figure 5.1. Reaction time (ms) for words for body N size by length for the native English speaker group……………………………………………………..170

Figure 5.2. Reaction time (ms) for words for body N size by length for the native German speaker group……………………………………………………..171

Figure 5.3. Reaction time (ms) for words for body N size by length for English participants (Ziegler et al., 2001)…………………………………………173

Figure 5.4. Reaction time (ms) for words for body N size by length for German participants (Ziegler et al., 2001)…………………………………………174

Figure 5.5. Reaction time (ms) for nonwords for body N size by length for the native English speaker group……………………………………………………….177

Figure 5.6. Reaction time (ms) for nonwords for body N size by length for the native German speaker group……………………………………………………….178

Figure 5.7. Reaction time (ms) for nonwords for body N size by length for English participants (Ziegler et al., 2001)…………………………………………179

Figure 5.8. Reaction time (ms) for nonwords for body N size by length for German participants (Ziegler et al., 2001)…………………………………………180

Figure 5.9. Accuracy for words for body N size by length for Native English speakers……………………………………………………………………...181

Figure 5.10. Accuracy for words for body N size by length for native German speakers……………………………………………………………………..182

Figure 5.11. Accuracy for nonwords for body N size by length for native English speakers…………………………………………………………………….186

Figure 5.12. Accuracy for nonwords for body N size by length for native German speakers……………………………………………………………………187

Figure 5.13. Length effect for nonwords in English in Experiment 3 and for German in Ziegler et al. (2001)…...…………………………………………………….188

Figure 5.14. Length effect for words in English in Experiment 3 and for German in Ziegler et al. (2001)……………………………………………………….....189

Figure 5.15. Length by body N interaction for native German speakers in Experiment 3……………………………………………………………….190

LIST OF TABLES

Table 2.1 Rule differences between English and German (Zeigler et al, 2000)…………36

Table 2.2 Example of Stimulus Materials for Experiment 1……………………………40

Table 2.3 Summary of Proposed Experiments and Tasks………………………………46

Table 3.1 Mean Language History Questionnaire data for Native German and Japanese Speaker Groups……………………………………………………………………58

Table 3.2 Reading Span results for native English and native German speakers………..60

Table 3.3 Mean RT (ms) and accuracy (out of 100%) by syllable length and nonword type for Experiment 1……………………………………………………………61

Table 3.4 Mean RT (ms) for native English speaker word data in Experiment 1……….66

Table 3.5 Mean accuracy for word data (out of 100%) in Experiment 1………………..67

Table 3.6 Nonword RT (ms) and accuracy data (out of 100%) from Experiment 1 and Goswami et al. (2001) collapsed across syllable length………………………..69

Table 3.7 Mean RT (ms) and accuracy (out of 100%) by syllable length and nonword type for Native German speakers for Experiment 1…………………………...72

Table 3.8 Mean RT (ms) for native German speaker word data in Experiment 1……….78

Table 3.9 Mean accuracy (out of 100%) for native German speaker word data in Experiment 1……………………………………………………………………79

Table 3.10 Mean RT (ms) by syllable length, nonword type, and proficiency for Native Japanese speakers for Experiment 1…………………………………………..81

Table 3.11 Mean accuracy (out of 100%) by syllable length and nonword type for Native Japanese speakers for Experiment 1…………………………………………..84

Table 3.12 Mean RT (ms) for native Japanese speaker word data in Experiment 1…….87

Table 3.13 Mean accuracy (out of 100%) for native Japanese speaker word data by syllable and proficiency rating in Experiment 1………………………………88

Table 3.14 Mean RT (ms) for syllable type by language group…………………………89

Table 3.15 Mean accuracy (out of 100%) for syllable type by language group…………91

Table 3.16 Mean RT (ms) for syllable type by language group for words………………94

Table 3.17 Mean accuracy (out of 100%) for syllable type by language group for words………………………………………………………………………….96

Table 3.18 Mean RT (ms) for nonword type by language group………………………..98

Table 3.19 Mean accuracy (out of 100%) for nonword type by language group………101

Table 4.1 Reaction time (ms) as a function of type of parse and number of phonemes for words in Experiment 2 and Treiman and Chafetz (1987)…………………….112

Table 4.2 Reaction time (ms) as a function of type of parse and number of phonemes for nonwords in Experiment 2 and Treiman and Chafetz (1987).………………..118

Table 4.3 Reaction time (ms) as a function of type of parse and number of phonemes for words by SOA in Experiment 2……………………………………………….123

Table 4.4 Reaction time (ms) as a function of type of parse and number of phonemes for nonwords by SOA in Experiment 2.…………………………………………..125

Table 4.5 Accuracy (out of 100%) as a function of type of parse and number of phonemes for words in Experiment 2 and Treiman and Chafetz (1987)…………………128

Table 4.6 Accuracy (out of 100%) as a function of type of parse and number of phonemes for nonwords in Experiment 2 and Treiman and Chafetz (1987)……………..132

Table 4.7 Accuracy (out of 100%) as a function of type of parse and number of phonemes for words by SOA in Experiment 2…………………………………………...137

Table 4.8 Accuracy (out of 100%) as a function of type of parse and number of phonemes for nonwords by SOA in Experiment 2……………………………………….140

Table 4.9 Reaction Time (ms) for native German speakers as a function of type of parse and number of phonemes for words in Experiment 2…………………………143

Table 4.10 Reaction time (ms) for native German speakers as a function of type of parse and number of phonemes for words by SOA in Experiment 2……………….147

Table 4.11 Reaction Time (ms) for native German speakers as a function of type of parse and number of phonemes for nonwords in Experiment 2…………………….149

Table 4.12 Reaction time (ms) for native German speakers as a function of type of parse and number of phonemes for nonwords by SOA in Experiment 2…………...150

Table 4.13 Accuracy (out of 100%) for native German speakers as a function of type of parse and number of phonemes for words in Experiment 2…………………152

Table 4.14 Accuracy (out of 100%) for native German speakers as a function of type of parse and number of phonemes for words by SOA in Experiment 2……….153

Table 4.15 Accuracy (out of 100%) for native German speakers as a function of type of parse and number of phonemes for nonwords in Experiment 2…………….154

Table 4.16 Accuracy (out of 100%) for native German speakers as a function of type of parse and number of phonemes for nonwords by SOA in Experiment 2…..156

Table 5.1 Reaction Time (ms) for words by body neighborhood size and word length for native English speakers, native German Speakers, and Ziegler et al. (2001)..169

Table 5.2 Reaction Time (ms) for nonwords by body neighborhood and word length for native English speakers, native German Speakers, and Ziegler et al. (2001)..176

Table 5.3 Accuracy (out of 100%) for words by body neighborhood size and word length for native English speakers and native German Speakers……………………183

Table 5.4 Accuracy (out of 100%) for nonwords by body neighborhood size and word length for native English speakers and native German Speakers…………….185

ACKNOWLEDGEMENTS

Although words cannot adequately describe the gratitude and appreciation I have for all who supported me through this process I would first like to thank my advisor Judy Kroll. Although it is impossible to mention all of the things that I am grateful for here on these pages I would like to thank her for welcoming me into her lab for my final years of graduate school and making the transition smooth and easy. I am also indebted to the countless hours and the overwhelming enthusiasm that she has had for this project and the other projects that we have worked on over these years. She is an incredible mentor and advisor and I will miss our weekly meetings and her boundless energy.

I would like to thank Lael Schooler for being a great advisor and mentor for the first part of my graduate career, for all of his input in the early stages of this project, and his continued support and friendship. I would also like to thank the members of my committee, Giuli Dussias, Rick Gilmore, and David Rosenbaum, for all of their insightful feedback on this project. It has been really wonderful working with all of you during these past years.

I am also indebted to the College of Liberal Arts for the generous dissertation grant which made it possible to travel to Germany to collect data. Without this grant it would have been difficult to recruit the number of German-English bilinguals necessary to complete this project.

I would also like to express my gratitude to those who made my trip to Germany a wonderful experience. I would like to thank Pienie Zwitserlood for opening up her lab to me and giving me a wonderful place to work during my data collecting visit to the

University of Muenster. I thank Jens and Heidi for making me feel welcome, showing me the ropes of subject recruitment, and introducing me to friends and future participants; and Christian for all of his help making arrangements for me before I arrived and his help with mustering subjects during the end of my stay in Muenster.

I thank Sonja Kotz for allowing me to visit her at the Max Planck Institute in Leipzig and for all of her feedback on this project and our chats about research and life while I was there. Her enthusiasm for my project made it possible for me to recruit many wonderful participants during my short stay.

I am extremely grateful to everyone who provided me a place to stay while I finished writing in State College after our move. I thank Shannon and Fred, Maya and Chris, and Silvia and April for generously opening up their homes to me this summer and providing me with quiet and comfortable places to write.

I would also like to thank everyone who helped implement this project. I am thankful for Eva Suarez for all of her hard work on the German version of the reading span task and for helping to recruit German participants and collecting data. I also cannot thank her enough for being my overall German expert. I would like to thank Katie Moran, Stephanie Rodriguez, and Ashley Yuhouse whose hard work and conscientious data collection helped make this project a success. Your dependability allowed everything to run smoothly while I was in Germany this past spring.

I would also like to thank everyone whose friendship made graduate school a much more pleasant experience and especially the women of the Purple Lab for their undying support through this process and for listening patiently to every update. I will greatly miss our time in the lab together! I would like to thank April Jacobs whose unique sense of humor always kept me laughing; Sara Hasson for her style and energy and most importantly her knack for knowing just what to say to PsyScope; Maya Misra for our chats about pseudohomophones, research, and life; Noriko Hoshino for all of our fun road trips, early morning chats in the lab and the overall goal that someday we had to get to the lab in the morning before Judy got to her office; and Ana Schwartz, my “twin” who survived comps and the dissertation with me. I am glad that we are finally done but I will greatly miss our coffee talk, sushi, and duck feeding adventures.

Finally, I would like to thank my family who has lived through this process along with me. To my parents who taught me to love learning early in life, to value education, and who have always supported my dreams. I thank my mother for always knowing the right words of encouragement to keep me going through this process and my dad who proudly kept track of my daily page count, which motivated me to write even more. I would also like to thank my mother- and father-in-law for their encouragement and my father-in-law for his advice as someone who has lived through this process. Last but not least, I would like to thank Brian, my husband, for making these last few years of graduate school fun and memorable and for weathering through this final part with me. I look forward to all that the future has in store for us.

Chapter 1: Background and Introduction

Anyone who has ever tried to become proficient in another language can attest that acquiring a second language as an adult is not an easy task. Cognitively, learning a second language may be difficult because the learner has to sort out and understand the similarities and differences that exist between the first (L1) and the second language (L2). These similarities and differences between languages can occur on many levels. For example, one of the more obvious differences between English and Japanese is that English uses a Roman alphabet and Japanese has a logographic orthography. But even for languages that share the same script, differences often can be found in the phonological representations and grammatical structures.

In addition to representational differences, languages can differ in the processing preferences used to retrieve and produce orthographic, phonological, semantic, and grammatical representations. Cross-linguistic research examines these differences by comparing the processing strategies of native speakers from different language groups. At the level of orthography, these studies provide a great deal of evidence that the consistency with which spelling is mapped onto the sounds of the language has an impact on how the representations are processed during word recognition and production (Goswami, Ziegler, Dalton, & Schneider, 2001; Katz & Feldman, 1983; Martensen, Maris, & Dijkstra, 2000; Ziegler, Perry, & Coltheart, 2000; Ziegler, Perry, Jacobs, and Braun, 2001). The results of these cross-linguistic studies, showing distinct preferences for different languages, might be taken to suggest when an individual is bilingual in languages that rely on different strategies, these processes themselves might serve to differentiate the two languages. However, recent bilingual research on speakers of languages that differ in their parsing preferences, demonstrates that even when a bilingual is using one language only, the other language is active (Dijkstra, van Jaarsveld & Ten Brinke, 1998; Dijkstra & van Heuven, 1998; Hermans, Bongaerts, De Bot, & Schreuder, 1998; Miller & Kroll, 2001; van Heuven, Dijkstra & Grainger, 1998). This apparent paradox raises the question of whether or not language-specific information is acquired and used by learners during second language processing. If proficient second language users do have access to this processing information, how do learners make sense of the similarities and differences between their first language and the language being learned? Furthermore, is the difficulty of learning a second language dependent on learning the representations of that language (i.e., words, sounds, and grammar), or learning new strategies to parse and analyze the new language? For example, do native speakers of German find it easy to learn the English word *ball* because it shares the same spelling, meaning and basic pronunciation as the German equivalent *ball*? Or is learning aided by the fact that like German, the English pronunciation of the word *ball* and other orthographically similar words (i.e. *call*, *fall*, and *tall*) is highly consistent so it does not matter whether *ball* is parsed by the learner as *b*/*all*, *ba*/*ll*, or *b*/*a*/*ll*?

One goal of the present research is to determine what lexical information and strategies learners use and develop while acquiring another language, and to uncover whether or not L2 learners eventually use the same lexical parsing strategies as native speakers of the language or whether the resulting strategies reflect a contribution from both of the bilingual's languages. More specifically the goal is to examine how bilinguals process L1 and L2 words during recognition and production and to identify the basic sublexical units, or parts of words, that are accessed in the first and second languages.

Cross-linguistic differences in language processing

Evidence from past cross-linguistic studies shows that native speakers across languages process their native languages in different ways at both the sentence and the word level. Although past research suggests that for bilinguals, both languages are active and ready to compete for selection (Dijkstra et al., 1998; Hermans et al, 1998; Jared & Kroll, 2001; Van Heuven et al., 1998), very few studies examine how the degree of nonselectivity may differ in bilinguals whose languages have different parsing preferences. Most current research focuses on which codes are active (Dijkstra et al., 1998; Dijkstra, Grainger, & Van Heuven, 1999; Van Heuven et al., 1998), some research focuses on parsing (Cutler, Mehler, Norris, & Segui, 1989; Pallier, Colomé, & Sebastián-Gallés; 2001), but none focuses on how parsing preferences may interact with nonselective access.

Evidence on syntax

Cross-linguistic studies of parsing. The majority of research providing evidence that processing varies across languages and within bilinguals has been conducted in the domain of syntax and phonology. One way in which processing varies across languages is the way that speakers across languages resolve ambiguity at the sentence level. Consider ambiguous sentence, “The reporter talked with the daughter of the actor who had just been arrested.” For this sentence, it is unclear whether it was the daughter or the actor who was recently arrested. According to the Late Closure principle (Frazier, 1979; 1987), a sentence parsing principle proposed to be universal across all languages, new information (e.g., who had just been arrested) should be attached to the clause that is currently being processed (e.g., of the actor) and therefore the parser should interpret the actor as the one who was arrested. Although Late Closure does appear to be a preferred processing strategy in English (Frazier & Rayner, 1982), and Italian (Vincenzi & Job, 1995), Mitchell and Cuetos (1991) found evidence that this strategy was not universal when examining sentence processing preferences in Spanish.

In the Mitchell and Cuetos (1991) study, native Spanish speakers were presented with ambiguous and unambiguous sentences similar to the example above. The critical items were nonambiguous sentences where the new information at the end of the sentence attached to the clause that is being processed, thus forcing attachment to the lower of the two final nouns (e.g., actor, in the above sentence). This attachment was forced by making the information in the final part of the sentence agree only with the final noun through the use of agreement principles in Spanish (e.g., gender agreement). The processing of these forced low attachment sentences were compared with the processing of ambiguous sentences where the final information could attach low to the final noun, or high to the next to last noun (e.g., daughter). Participants processed these ambiguous sentences more quickly than those that forced low attachment, suggesting that native speakers of Spanish generally prefer a high attachment processing strategy and that the Late Closure principle proposed by Frazier (1979) is not universal.

The Competition Model. A model that has been used to examine cross-linguistic and also bilingual comprehension differences in syntactic processing is the Competition Model (for a review see MacWhinney, 1997). According to this model, language learners acquire syntax by learning the cues of the language. For example, in English the word order of a sentence is a much stronger cue as to what is the subject or agent of the sentence than agreement or animacy. In languages like Spanish and German that allow prepositions or articles to mark the subject or the object in a sentence, word order becomes more arbitrary and a less strong cue for comprehension. The competition model proposes that individuals learn cues through the environmental language input. How quickly they learn the cues of their language depends on how available those cues are in the linguistic environment of the learner, how frequently the cues appear in that environment, how reliable a cue is, and how often one cue comes into competition with another cue. Within the model, those cues that produce a reliable interpretation of a sentence are learned early and strongly guide the processing of native language processing in adults.

Empirically, there is a good deal of evidence that those cues that are reliable are acquired early. In a cross-linguistic study looking at children learning English, Italian, and Hungarian (Bates & MacWhinney, 1989), word order seemed to be acquired the earliest and the fastest for children learning English, but not for children learning Italian and Hungarian, two languages in which word order is more arbitrary. In contrast, case marking appeared to be the most reliable cue for children learning Hungarian, and for Italian children there was an early dependence on animacy but later in development, there was a greater reliance on agreement. Although this last example seems to contradict the Competition Model, it is not completely contrary since agreement is a difficult grammatical cue and it is possible that the younger children were not cognitively ready to use this cue as a dominant one until later in language development.

The Competition Model has also been used to try to predict sentence comprehension performance in second language learners. In the bilingual version of this model, the late language learner begins learning with a highly developed and efficient first language system already in place. At the beginning of the acquisition process, the learner uses sentence processing strategies from the first language until the cues from the second language can be acquired. Sometimes syntax transfers easily from the first language to the second language. For example the English sentence, "The student took the exam." translates into Spanish as "*El estudiente tomó el examen.*" This is an example of good transfer between Spanish and English as both sentences have not only have the same syntactic structure, but they also share cognates (student-*estudiente*, exam-*examen*). In contrast, there are many cases where the syntactic structure of an English sentence does not transfer to Spanish. For example, if you wanted to declare that "Marc sang." in Spanish, you would say, "*Cantó Marc.*" placing the verb before the subject.

The competition model predicts that the learner will first rely on and overuse the sentence processing strategies from the first language. Therefore, early language learners would have no problem translating the first example sentence from English to Spanish. However for the second sentence, learners should be less likely to use or take longer to comprehend this declarative structure and may produce, "*Marc cantó.*" Studies looking at sentence comprehension do find support for the Competition Model and show that when reading in their second language, learners often will interpret sentences using syntactic cues from their first language. For example, a study by McDonald (1987) found that over time as native English speakers learned Dutch, they began to rely less on word order during comprehension and more on the case inflection system of Dutch. This study also found the reverse pattern for Dutch native speakers learning English. As Dutch native speakers became more proficient in English, they began to rely more on word order.

Furthermore, research examining the syntactic level of processing in bilinguals indicates that parsing strategies from the second language are sometimes transferred to the first language if those parsing strategies help decrease the demands on memory. In a study by Dussias (2001), the sentence parsing strategies of Spanish-English and English-Spanish bilinguals were examined using ambiguous sentences such as *John visited the sister of the actress that lived in Bel Air*. These two bilinguals groups were of interest because as mentioned earlier, monolingual Spanish speakers generally resolve the ambiguity of sentences like the one above through the use of high attachment (Cuetos & Mitchell, 1988; Mitchell & Cuetos, 1991). In other words, Spanish speakers tend to attach the ambiguous phrase "*that lived in Bel Air*" with "*sister*," the first or "higher" of the two nouns in the complex noun phrase. In contrast, English monolinguals tend to resolve ambiguity through low attachment (Frazier, 1979; 1987), attributing the ambiguous phrase to "*actress*," the second or "lower" or the two nouns. For this study Spanish-English and English-Spanish bilinguals were presented with ambiguous sentences in Spanish and in English and then a comprehension question. For example, participants may have seen a sentence like the example presented above: *John visited the sister of the actress that lived in Bel Air*. Participants would then be given the question *Who lived in Bel Air?* along with the choice of two statements: *The sister lived in Bel Air. The actress lived in Bel Air.* The performance of the two bilingual groups was then compared with monolingual speakers of English and Spanish. The surprising finding for this study was that the Spanish-English bilinguals had not only learned to use a low attachment strategy when processing English, but also transferred this strategy back to their L1.

Evidence on lexical transfer

The studies reviewed above demonstrate that second language syntactic strategies can be acquired by learners and that sometimes second language strategies can be transferred to the first language if the strategy decreases cognitive load during processing. The Competition Model addresses and makes clear predictions about how sentence level cues transfer from the first language to the second language and compete with the new second language cues, however it is not clear how the mechanisms within the model can account for the transfer of L2 sentence processing strategies on the L1.

Although the Competition Model makes clear predictions about transfer at the syntactic level, it has not been used to examine how processing strategies at the word level may transfer and compete in the bilingual. One model that does address how processing may change over time at the word level is the Revised Hierarchical Model (Kroll & Stewart, 1994). In this model, the second language learner has a word store for each language and a shared semantic store (see Figure 1.1). According to the Revised Hierarchical Model, at the early stages of acquisition, the learner is highly dependent on the first language when trying to retrieve the meanings of words in the second language. This dependence on the first language for access to meaning is demonstrated by the strong lexical link from words in the first language to words in the second language. As learners become more proficient, they are eventually able to conceptually mediate their second language. In other words, they are able to retrieve meanings for second language words directly from the conceptual store.

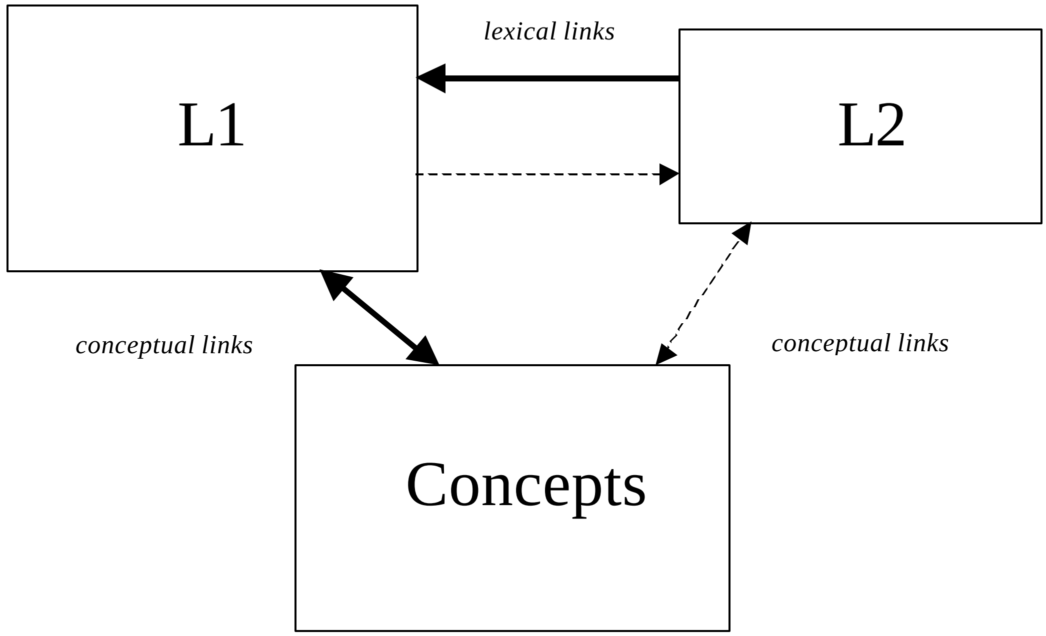


Figure 1.1. The Revised Hierarchical Model (adapted from Kroll & Stewart, 1994).

There have been a number of studies supporting the claims of the Revised Hierarchical Model (Dufour & Kroll, 1995; Dufour, Kroll, & Sholl, 1996; Kroll & Stewart, 1994; Sholl, Sankaranarayanan, & Kroll, 1995). Talamas, Kroll, and Dufour (1999) reported a study that clearly illustrates a strategy shift from less to more proficient bilinguals. More and less proficient English-Spanish bilinguals performed a translation recognition task in which they were presented with Spanish and English word pairs and had to decide whether the pairs were translation equivalents. For example, if a participant saw the pair *hombre*-man, the response given would be "yes." The distractor trials in this task were manipulated such that the incorrect Spanish word was related in form or meaning to the correct translation. For example, a form distractor for the correct pair presented above could be the Spanish word *hambre* which means “hunger.” A meaning related distractor would be the word *mujer* which is Spanish for “woman.”

Talamas et al. (1999) found greater interference for the form- and meaning-related distractor items than for nonrelated control items. More importantly, the two proficiency groups differed as predicted by the Revised Hierarchical Model. According to the model, to determine whether two words are translation equivalents, the less proficient bilinguals should be more reliant on translating the L2 form to the L1 form to retrieve its meaning. Therefore, as the model predicts, this group had more interference from the form distractors than the meaning distractors. The model also predicts that the more proficient group should be able to directly access meaning from the L2 word and should rely less on the word form. As predicted, this group had greater interference from the meaning related distractors than the form related distractors, demonstrating a strategy shift from form to meaning with increased proficiency.

The general evidence presented above concerning syntax and meaning in second language processing shows that the learner first is dependent on L1 forms and strategies and then slowly becomes able to process L2 more independently. At the level of semantics, concepts for L2 words seem first to be retrieved through the translation equivalent from L1. For syntax, during acquisition the learner transfers grammatical constructions from their first language to the second language until the grammatical constructions from the second language become strong enough to be used (e.g., MacWhinney, 1997). Although some work has been done looking at the transfer of processing strategies from the first language to the second language at the syntactic level, there has been less research examining the effect of processing transfer and the constraints of transfer at the lexical level.

Constraints on parsing: evidence from phonology

The few studies that have investigated distinct processing preferences across languages focus on what could be considered the limits of sensitivity to word-level cues in the second language. For example, a study by Cutler et al. (1989) examined auditory parsing strategies in proficient French-English and English-French bilinguals. The comparison between these two languages is interesting because French has very clear syllable boundaries unlike English whose syllable boundaries are often ambiguous. When asked to respond to a specific phonemic segment, Cutler et al. found that those bilinguals dominant in French parsed spoken French words like French monolinguals and English words like English monolinguals. That is, French dominant bilinguals were faster to detect the segment when it was an entire syllable than when it was not. However English dominant bilinguals did not show this parsing preference when detecting phonemic segments in French. The results suggested that the French dominant bilinguals were able to use the syllable parsing strategy when listening to a language with clear syllable boundaries (French) and were also able to “turn off” that strategy when listening to words in English, a language with less clear syllable boundaries. The English dominant participants, although highly proficient in French, were never able to acquire the syllable parsing strategy.

A more recent study by Pallier et al. (2001) investigated auditory processing of Spanish-Catalan bilinguals and Catalan-Spanish bilinguals listening to Catalan words.

Participants were presented with an auditory prime in Catalan and then were asked to make a lexical decision to the target word. Participants made a lexical decision by pressing one button if the target stimulus was a word and another button if the stimulus was a nonword. The critical prime-target stimulus pairs for this study were those that contained a phonological feature that was distinctive in Catalan but not in Spanish. The ability to process this distinctive feature made these stimulus pairs homographs, words similar in form but distinct in meaning. Therefore, no priming should occur for these pairs. The inability to process this feature made the pair homophones, words identical or nearly identical in spelling, and pronunciation (e.g., més-*mes*; more-*month*). Failure to process this feature should result in a repetition priming benefit for the target word. Pallier et al. found that Catalan-Spanish bilinguals did not show priming for the homograph pairs containing the distinctive feature, they only showed priming for identical prime-target pairs. However the Spanish-Catalan bilinguals, showed the same amount of priming for the distinctive feature homograph pairs as they did with homophone prime-target pairs. This finding suggests that the Spanish-Catalan bilinguals were unable to process this distinctive feature online and therefore were processing Catalan using the phonological constraints of their first language, Spanish.

The studies by Pallier et al. (2001) and Cutler et al. (1989) demonstrate that different processing strategies across languages can affect and constrain bilingual word recognition. Although these studies were conducted in the auditory domain with bilinguals, the demonstration of different processing strategies across languages is very much in line with cross-linguistic studies looking at the processing of native speakers in two or more language groups. This evidence will be presented in the following section.

Evidence for unique cross-linguistic lexical processing

Many past studies examining English word recognition have yielded clues about how native speakers organize and retrieve words from memory (for a review see Andrews, 1997). Because English is an orthographically deep language, meaning that there are multiple mappings between spelling and sound, there have been a number of studies that have tried to examine the critical units involved in processing English. A number of recent findings have suggested that the orthographic rime, the first vowel plus the remaining letters in a one-syllable word, may be an important unit of processing in English. For example the orthographic rime for the word *bat* would be *–at*. A series of statistical analyses of the English language by Treiman and her colleagues found that the pronunciation of a vowel in English is better predicted by the letters that follow the vowel than the letters that precede it and that the cluster of the vowel plus the following letter or orthographic rime, is a better predictor of the pronunciation of the vowel than the vowel alone (Treiman, Mullennix, Bijelac-Babic, Richmond-Welty, 1995; Kessler & Treiman, 1997). For example, in English the letter “a” can be pronounced a number of different ways (e.g., bat, base, and ball). However knowing that the only letter following the letter “a” is a “t” (forming the rime –at) is much more helpful than knowing that the letter preceding the “a” is a “b” because the pronunciation of “a” in the rime unit –at is statistically more consistent than the pronunciation of a in the letter cluster “ba-.”

Additional support for the importance of the orthographic rime in English word recognition comes from an experiment conducted by Ziegler and Perry (1998) who suggested that this processing unit may explain differences in neighborhood effects across languages. In English, word recognition is facilitated when a given word has a large number of neighbors, words that differ from the target word by only one letter. However, in languages in which there is a more consistent spelling-to-sound mapping such as French and Spanish, there are null or inhibitory effects for neighborhood size (i.e., Carreiras, Perea, & Grainger, 1997; Johnson & Pugh, 1994). To examine whether the cross-linguistic difference for neighborhood size is a result of the consistency of the rime unit in English, Ziegler and Perry compared English words with a large number of rime neighbors (neighbors that share the same orthographic rime) to words with a small number of rime neighbors while keeping traditional neighborhood size constant. In a second manipulation, the number of rime neighbors was held constant while the number of traditional neighbors varied. The results of this experiment showed facilitation for large numbers of rime neighbors and null effects for traditional neighbors. In addition, they provide converging evidence for Treiman's (Kessler & Treiman, 1997; Treiman & Chafetz, 1987; Treiman et al., 1995) findings that the rime is important for reading and pronouncing English.

Although there is considerable evidence that the orthographic rime is an important processing unit in English, cross-linguistic evidence suggests that this unit may only be important to languages like English that do not have consistent spelling-to-sound correspondences. In a statistical examination of Dutch similar to Treiman et al. (1995), Martensen et al. (2000) found that unlike English, the rime unit did not predict naming latencies of Dutch words. Instead, smaller sublexical units such as the onset, nucleus, and coda were much better predictors. Martensen et al. attributed this difference to the consistency of Dutch. In other words because Dutch has a much more consistent spelling-to-sound mapping, the pronunciation of "a" can be predicted just as well given the preceding letter “k” as it can from the following letter “t,” making the information provided by the rime unit redundant. Therefore, unlike English, Dutch is a more consistent language and smaller units can be used to parse words.

The Dual Route Model. Other cross-linguistic studies provide additional evidence for the idea that differences in orthographic depth across languages correspond to differences in lexical parsing. For example, Ziegler and his colleagues (Ziegler, Perry, & Coltheart, 2000) extended the Dual-Route Cascaded model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) to compare the grapheme-to-phoneme correspondence rules in English and German. The Dual-Route Cascaded Model (Coltheart et al., 2001) has computationally implemented two routes for word recognition and naming. Two routes were included in this model to capture both parallel and serial processing aspects of word recognition and production. In the lexical nonsemantic route, phonology is directly retrieved, and in the grapheme-to-phoneme correspondence, or GPC route, phonology is serially assembled letter by letter (see Figure 1.2).

The lexical nonsemantic route is based on McClelland and Rumelhart’s (1981) Interactive Activation Model. When a word is presented to the model, the model first extracts the features of the word which then spread activation to stored letters and then the activated letters in turn spread activation to the orthographic and phonological representations of that word.

The GPC route shares the feature level and the letter level with the lexical nonsemantic route, however in this route rules are applied serially to each letter until the phonology of a given word is computed. The rules used by the GPC route are derived through an algorithm that looks at the letter-to-sound mappings of all the words in the model's lexicon.

The Dual-Route model has had success in accounting for a large number of results in lexical decision and word naming tasks within a single language including frequency effects, the general finding that words that appear more frequently in print are able to be recognized as a word and named faster than words that appear less frequently in print (Andrews, 1989; Scarborough, Cortese, & Scarborough, 1977);

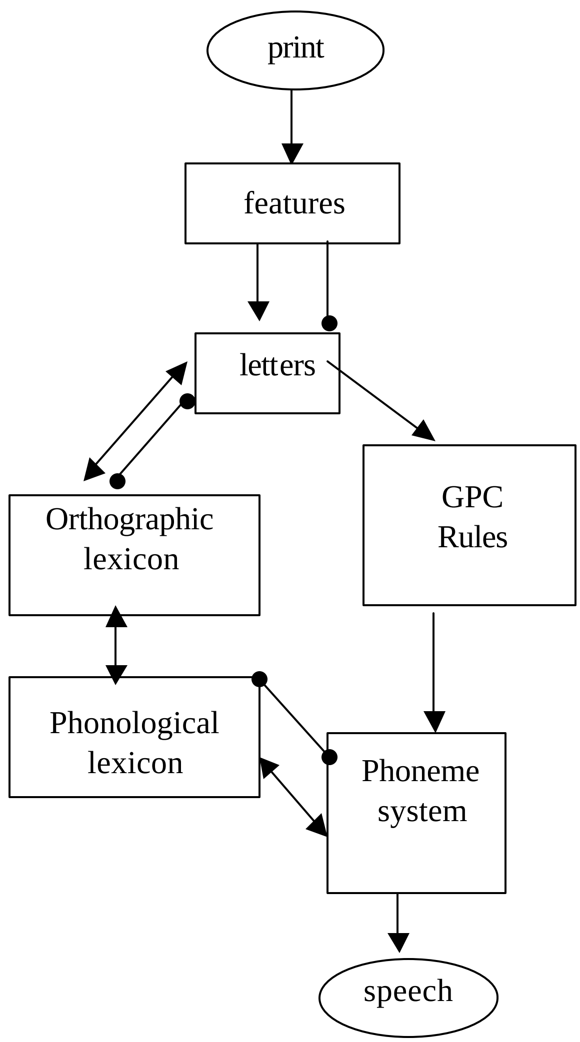


Figure 1.2. The DRC Model (adapted from Coltheart et al., 2001).

and pseudohomophone effects, nonwords that can be pronounced like real words in English (e.g., kart) are pronounced faster than nonpseudohomophonic nonwords and take longer to reject in a lexical decision task (Coltheart, Davelaar, Jonasson, & Besner, 1977; Rubenstein, Lewis, & Rubenstein, 1971; see Coltheart et al., 2001 for a complete review). Despite its success, there has been a long standing debate over whether having two routes is necessary since models such as the Parallel Distributed Processing model (Harm & Seidenberg, 1999; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) are also able to simulate a wide variety of word recognition and production tasks without the need for two routes. However the debate over which model best represents word recognition and production in normal and impaired language processing is beyond the scope of this dissertation.

Extending the Dual Route Model to German. Because of the success of the Dual Route model in English, Ziegler et al. (2000) extended the model to German, using the same algorithm to extract the grapheme-to-phoneme correspondences for German, an orthographically deep but consistent language. First the model was given a German vocabulary. To determine the GPC rules, the words of the model's vocabulary were segmented into their respective graphemes and phonemes. Next the same algorithm from the English model was used to extract the grapheme-to-phoneme correspondences for German (Ziegler et al., 2000). After using the model to correctly name a list of German words, the model was then tested to determine whether it could produce the German loan word effect, which is somewhat equivalent to regularity effects in English. Like humans, the model demonstrated that loan words (German words borrowed from other languages) that violated German grapheme-to-phoneme correspondence rules were named more slowly than words that only violated German orthography which in turn were named more slowly than words conforming to German orthography. The successful simulation of this finding in German suggests that the Dual-Route architecture can be extended to other languages besides English and that a two- route model may be appropriate for languages that are more consistent than English. In addition, this extension also demonstrates that while the overall two- route architecture can be extended across languages, the computational rules used by the assembled route do differ from language to language and that differences in orthographic depth across languages do have consequences on processing.

Cross-language differences in parsing strategies in native German and English speakers also have been examined using English-German cognates. Ziegler, Perry, Jacobs, and Braun (2001) examined rime effects and length effects on a list of cognates and nonwords that were nearly identical between both languages. In this experiment words and nonwords differed in length from 3 to 6 letters and also in the number of rime neighbors. The results of this experiment showed that native English speakers exhibited stronger effects for the number of rime neighbors than native German speakers. Conversely, native German speakers showed a stronger word length effect than native English speakers overall suggesting that native German speakers are using smaller units of processing than native English speakers.

A similar conclusion was reached in a developmental study comparing strategies used by English and German children learning to read (Goswami et al., 2001). The critical stimuli for comparing parsing strategies were pseudohomophones, nonwords that sound like real words (e.g., kake). In two experiments, 7-9 year old children were first asked to name pseudohomophones and then perform lexical decision on a list of words containing pseudohomophones, non-pseudohomophonic nonwords (i.e., dake), and real words. The results showed that for naming, English children made significantly more errors in naming pseudohomophones than German children, however for lexical decision, German children made significantly more errors than their English monolingual counterparts. The different pattern of errors for the two languages suggests that because German has a more consistent spelling-to-sound correspondence than English, German speakers rely more heavily on these correspondences than speakers of English.

The studies reviewed above suggest that there are distinct lexical parsing strategies associated with languages that differ in orthographic depth and consistency. However, few studies have asked what the implications of these differences might be for bilinguals. Do bilinguals transfer the parsing strategy associated with their L1 to the L2? If they are able to acquire parsing strategies that reflect the distinct lexical structure of each of their languages, do those differences limit the extent to which lexical transfer occurs across languages?

Cross-language interactions in bilingual word recognition

How different lexical processing preferences across a bilingual's two languages impact word recognition and production has largely been ignored in bilingual research for processing at the level of the word. Research on lexical access in bilinguals has focused on whether or not bilinguals have selective access to a language, or the ability to "turn off" one language when the other is being used; or whether lexical access is nonselective. The nonselective hypothesis assumes that both languages of the bilingual are active during recognition and production. Although some early evidence suggested that lexical access in bilinguals was selective (Gerard & Scarborough, 1989), more recent bilingual research has demonstrated that when a bilingual is recognizing or producing words in his or her second language, the first language is active to some degree (Dijkstra et al., 1998; Van Heuven et al., 1998).

Word recognition studies have demonstrated that the first language is active by showing that cognates, words that are identical or similar in form, pronunciation, and meaning are faster to recognize than matched L2-only controls (Dijkstra et al., 1998). In addition studies have showed that for L2 lexical decision, participants take longer to accept cross-language homographs, words that are similar in form and pronunciation but not meaning, as L2 words when the task requires them to make a no response to L1 words. However during a generalized lexical decision task, where participants are instructed to make "yes" responses to words in both the first and the second language, these same homograph items are responded to faster than matched controls (Dijkstra et al., 1998).

For production, much of the evidence for language nonselectivity has been found using picture-word interference (Hermans et al., 1998, but see Costa, Miozzo, & Caramazza, 1999 for an alternate explanation) and translation (Miller & Kroll, 2001). In the picture-word interference paradigm, pictures and words are presented and participants are instructed to name the picture or in certain cases name the distractor word that is either presented with or after the picture. Distractor words in these experiments generally vary in their degree of similarity to the L1 and L2 names for the picture. Results from studies using this paradigm show that when participants are asked to name pictures in L2, the presentation of a distractor word that is similar phonologically or semantically to the L1 name of the picture delays naming compared to an unrelated distractor (Hermans et al., 1998).

Findings similar to those of Hermans et al. (1998) have been found using a translation interference paradigm. In this paradigm, participants are asked to translate from one language to the other and are presented with distractors that vary in similarity to the L1 word and the L2 word (La Heij et al., 1990). A study by Miller and Kroll (2001) demonstrated that distractors that were similar in form to the required translation helped facilitate translation compared with unrelated distractors, again showing that information from both languages is active to some degree over the time course of processing.

In addition to trying to determine how active the two languages of the bilingual are during word recognition and production, another focus that has emerged in second language acquisition, bilingual word recognition, and production is generally on how the orthographic, phonological and syntactic representations or codes from L1 transfer (or fail to transfer) to L2. The majority of these studies have examined the overall degree of orthographic, phonological, and semantic overlap between words in both languages and have found that second language words that have a high degree of orthographic, phonological, and semantic overlap with words in the first language (i.e., cognates) are acquired more rapidly (de Groot & Keijzer, 2000; Lotto & de Groot, 1998) and also are more easily recognized and produced than second language words that have less overlap (Dijkstra et al., 1999; Dijkstra et al., 1998; Van Heuven et al., 1998). However, very few of these studies have taken into account the possible word parsing strategies imposed by the basic lexical structure of each language.

The BIA Model. Current computational models of bilingual word recognition also assume that the sublexical units and the processing preferences for the two languages of the bilingual are the same. Models such as the Bilingual Interactive Activation (BIA) (see Figure 1.3) model operate under the assumption that words first activate a stored set of perceptual features. These features in turn activate the letters that contain those features and then letters activate words that contain those letters and inhibit those that do not.

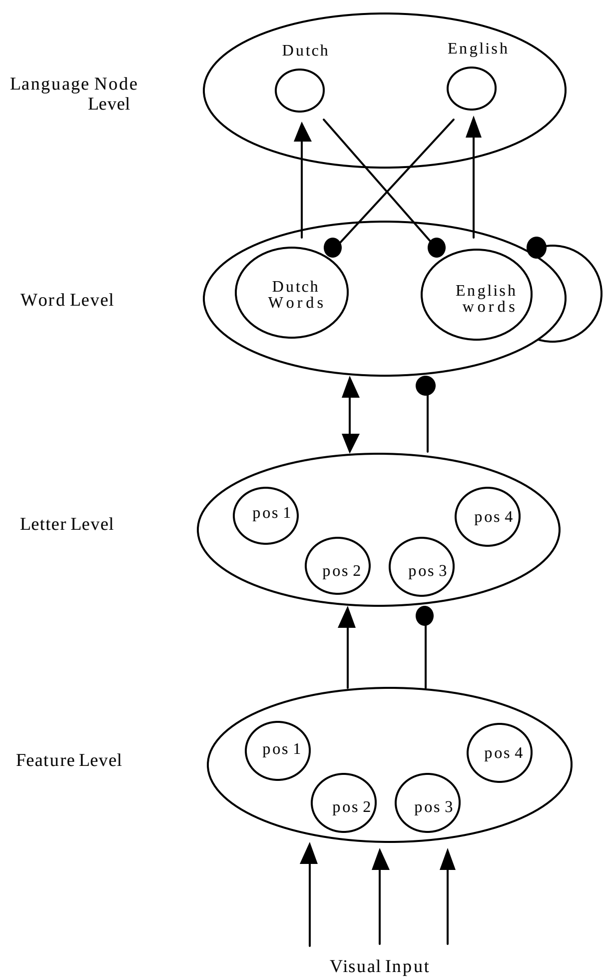


Figure 1.3. The BIA model of bilingual word recognition (adapted from Dijkstra & van Heuven, 1998).

Selection occurs as activation spreads to the language node providing top-down inhibition to words from the inappropriate language (Dijkstra et al., 1998). In localist models such as the BIA, because activation spreads directly from features to letters to words, they leave out the possibility that there may be mediating sublexical units between the letter and word level arising from the structure of the language.

Despite this possible shortcoming, the BIA model has been successful in capturing a number of findings from the bilingual word recognition literature. It is able to simulate the influence of word neighbors within the target language and nontarget language (Van Heuven et al., 1998). Using progressive demasking and lexical decision with Dutch-English bilinguals separately in English and Dutch, Van Heuven et al. found behavioral evidence that Dutch target words with large numbers Dutch neighbors were responded to more slowly than words with few Dutch neighbors. For English target words the pattern was reversed with shorter latencies for English target words with large numbers of English neighbors. However English target words with many Dutch neighbors were also responded to more slowly than English words with few Dutch neighbors.

The BIA model is also able to simulate effects of including cognates and interlingual homographs during lexical decision (Dijkstra et al., 1998). In a series of experiments, Dijkstra and colleagues found that when Dutch-English bilinguals were making lexical decisions in English, cognates were responded to more quickly than matched controls. However using the same task, in the next experiment they found that interlingual homographs such as *room* (meaning *cream* in Dutch), were responded to more slowly than matched controls when Dutch words were included in the list of nonwords requiring no responses. In the third experiment when the task was switched to generalized lexical decision, requiring yes responses for both English and Dutch words, interlingual homographs had shorter response times than matched controls. The BIA model was able to simulate the findings from these three experiments as well. However despite the model's success, BIA's underdeveloped semantic system and lack of a phonological system lead to the development of BIA+ (Dijkstra and van Heuven, 2002).

The BIA+ Model. The more recent version of BIA, BIA+ (see Figure 1.4) expands the original model by including sublexical orthographic and phonological units.

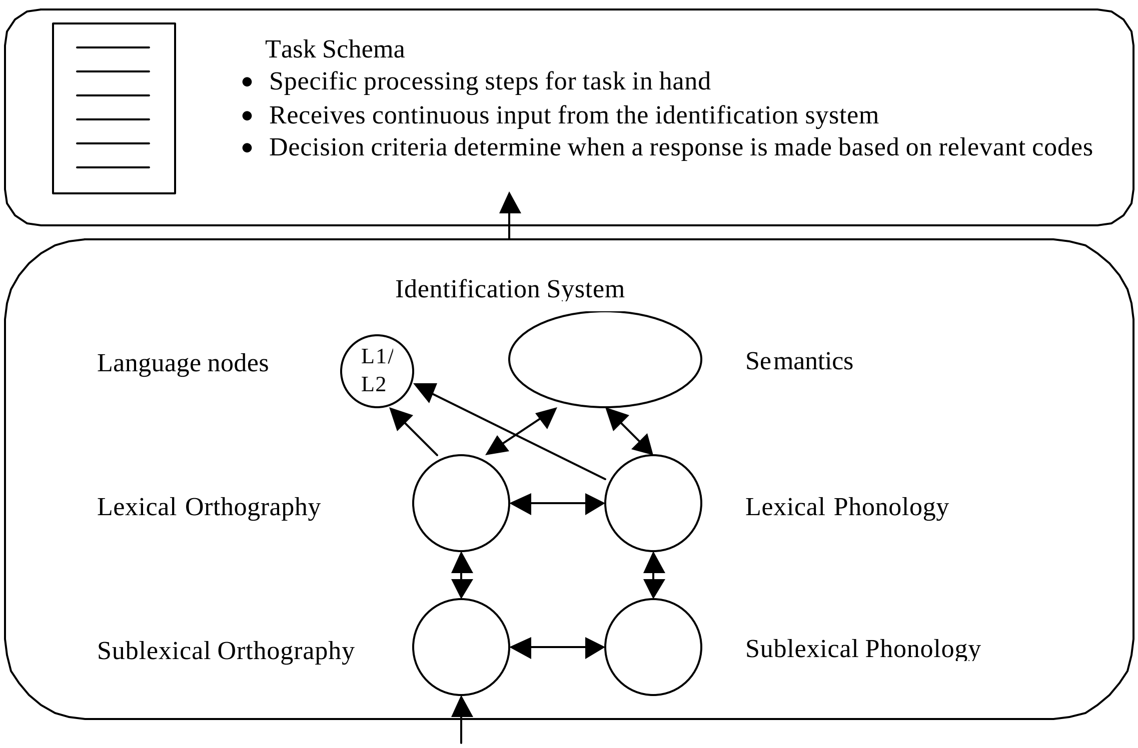


Figure 1.4. The BIA+ model of bilingual word recognition (adapted from Dijkstra & van Heuven, 2002).

The sublexical units represented in BIA+ include the onset (first consonant or consonant cluster), the nucleus (the first vowel or cluster of vowels), and coda (the remaining letters). One possible limitation of this model is that the same sublexical units are used when processing in either language within the model, suggesting that the same units are activated using the same bottom-up process for both languages (Van Heuven, 2000). However, it is not clear whether these units are the same across both languages, leaving the possibility that the strategies for parsing words and the mediating sublexical units vary across languages and that these different units help resolve the competition between two languages.

In addition, both the BIA and BIA+ generally characterize the performance of proficient bilinguals and have generally not been used to make developmental predictions about initial L2 acquisition. In one study, however, the BIA model was able to simulate performance for both more and less proficient bilinguals from a masked orthographic priming study (Dijkstra, Van Heuven, & Grainger, 1998). For this task, the performance for the less proficient French-English bilinguals was simulated first by lowering the relative L2 word frequency compared to the more proficient bilinguals, and also by lowering the amount of cross-language inhibition for the less proficient L2 speakers.

Although Dijkstra and his colleagues have addressed how the BIA models could be extended to look at acquisition (Dijkstra et al., 1998), these models do not make assumptions about any changes in the nature of representation that could occur over the course of language acquisition. Within its set of representations, the BIA+ model has only the sublexical units of onset, nucleus and coda. Although these units have been used successfully to simulate recognition and production in Dutch-English bilinguals, this set of representations excludes larger units that appear to be important to English word recognition such as the rime.

One addition that BIA+ includes that could allow the model to simulate parsing preferences is the task schema level. According to Dijkstra and van Heuven (2002), the task schema level has three main functions: carrying out specific processing relevant to the task, receiving input from the bottom-up identification system, and making an appropriate response when a set of criteria is reached. Although it is not completely clear how the BIA+ model would address the possibility of having distinct parsing preferences across the bilingual's two languages, it is possible that the task schema level could be a mechanism by which language-specific parsing preferences are stored and implemented. In addition, having the processing preferences at the task schema level also makes the clear prediction that the locus of these lexical parsing processes is quite late in the processing stream.

The Paradox

Although there is a great deal of evidence that processing preferences across languages may be different, research on bilingual word recognition and production has not considered the consequences of having two languages with different parsing preferences within one and the same individual. This is somewhat surprising considering that in many of these cross-linguistic studies, the native speakers are also proficient in more than one language. Instead, the bilingual studies have tried to determine how active words in one language are when the bilingual is recognizing or producing words in the other language. The cross-linguistic studies contrast parsing strategies across different languages in native speakers of those languages.

The goal of the present study was to take the methods from cross-linguistic studies that revealed different parsing preferences across language groups and use them to assess the performance of proficient bilinguals and second language learners. The specific tasks and how current models of word recognition may handle these tasks will be addressed in the following chapters. Some of the questions to be examined are: 1. Do nonnative speakers of English utilize different parsing preferences that yield different units of processing when recognizing and producing English? 2. Do these strategies vary depending on the script and consistency of the grapheme to phoneme mapping of the native language? and 3. If these strategies and units of processing are different, then how do processing units compete during recognition and production? Determining the basic components of processing for a given language as well as the statistics of how often these units occur will enable an understanding of how these units and processes compete during recognition and production. In addition, knowing which processing preferences are easier and harder to acquire may aid second language learners.

Chapter 2: Methods and Models

To address the questions raised in Chapter 1, a three-pronged approach was used employing tasks and methods from the cross-linguistic studies of Goswami et al. (2001), Ziegler et al. (2001), and Treiman and Chafetz's (1987) study looking at onset-rime parsing in native speakers of English. These tasks and methods were used to compare the performance of native speakers of English with proficient bilingual participants for whom English is the L2. The three tasks comprise a converging set of methods for examining lexical parsing.

In the first experiment, the critical task was lexical decision with pseudohomophones as the critical stimuli. This approach was used to examine the processing of pseudohomophones in a lexical decision task because these nonwords have been found to be a good indicator of how available phonology is during lexical decision and they are also a cross-linguistic measure of how efficiently spelling is mapped onto sound. In the second experiment, a modified lexical decision task with words and nonwords was used. This task, based on the work of Treiman and Chafetz (1987), violated hypothesized parsing preferences by inserting spatial and or temporal parsing guides and to observe the consequences for native and non-native speakers of English. In this task, the stimuli were revealed in stages with either the onset first and rime second, or with the onset and first vowel with the remaining letters second. The task in the third and final experiment was word naming with English-German cognates as the critical stimuli. This last approach was used to determine whether the cognate pairs in English and German that Ziegler et al. found to be named differently by native English and native German speakers are also named differently in the two languages by German-English bilinguals. More importantly this third experiment will examine whether L2 speakers of English are able to use L2 parsing strategies when the phonology is retrieved and also in the presence of strong L1 cues. This chapter will describe the logic underlying the comparison of these three tasks as well as outline experimental predictions based on the Dual-Route Model and BIA+ models.

A Tale of Two Models

The Dual-Route Model and BIA+ have been used to simulate a variety of language processing tasks and each model carries a set of assumptions that make unique predictions about the processing preferences used when reading words. Each model also has strengths and weaknesses for looking at the tasks and groups utilized in this set of experiments.

The Dual-Route Cascaded Model (DRC). The Dual-Route Cascaded Model has been successful in simulating a large number of findings for English word recognition and production. As mentioned in Chapter 1, this model processes words through a direct route and an assembled route. For the direct route all processing occurs in parallel. Letters activate orthographic representations of words which in turn activate phonological representations of words and then actual phonemes. It is this route that also contains lexical information such as word frequency. In contrast, processing in the assembled route is serial and cascaded. On each cycle of the model, letters activate corresponding phonemes and the phonology of a word is put together piece by piece. Because this route contains the grapheme-to-phoneme correspondence (GPC) rules for the language, this is the route that contains language specific processing strategies. By changing the assembled route to reflect the grapheme-to-phoneme consistencies of other languages, this model has also been extended to German (Ziegler et al. 2000) and more recently French (French DRC, n.d.). However the main weakness of this model is that it has not been extended to simulate word recognition and production in bilinguals. For example, in the German version of this model, only the rule set for the GPC route was changed, leaving the weighting of use for the two routes the same for English and German. In order to make predictions about how German-English bilinguals may process English stimuli, a set of assumptions will be made about what a bilingual version of this model might look like.

There are a number of ways that the Dual-Route Model could acquire a second language. The simplest way would be for the learner to create a second independent system with a lexical route and assembled route containing only the words, phonology, and grapheme to phoneme correspondence rules of the second language (Figure 2.1). However, the evidence demonstrating that both languages are active and competing during word recognition and production suggests that the bilingual would not possess these two independent systems.

A second way that the Dual-Route model could become bilingual is to have second language words only become part of the lexical route as learners acquire a second language (Figure 2.2). In this version of the model, the direct route would mostly resemble the BIA model of bilingual word recognition. Letters would activate orthographic representations of L1 words and L2 words and these representations would in turn activate stored phonological representations for the corresponding words. The GPC route would not be changed through acquiring the second language and rules for processing words in the first language would be used to assemble the phonology for words in the second language.

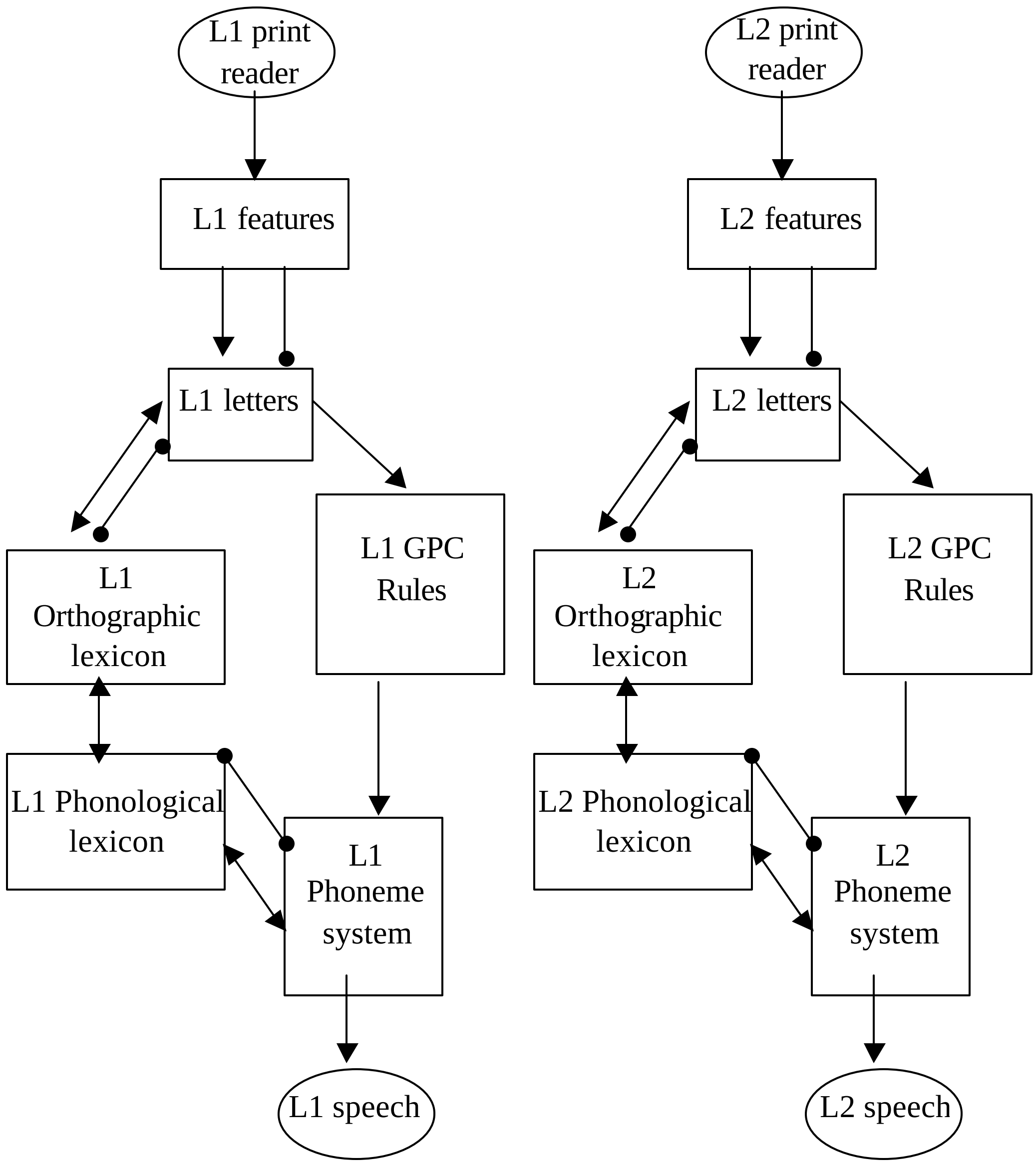


Figure 2.1. A bilingual DRC model with two independent processing systems for L1 and L2.

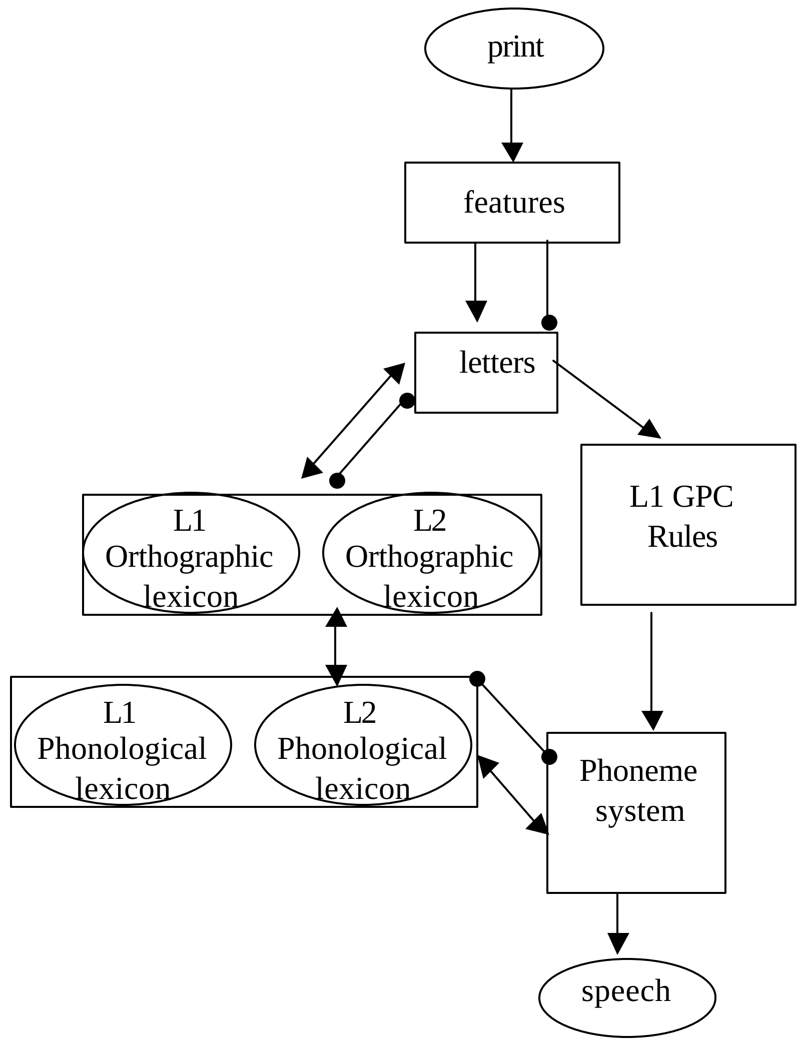


Figure 2.2. A bilingual DRC Model with L2 representations in the direct route.

A third possibility for a bilingual version of the Dual-Route model would be for second language information to be integrated within both routes of the model (Figure 2.3). In this bilingual version, in addition to second language lexical entries being added to the direct route, as a bilingual became more proficient in the second language, the GPC rules of that language would eventually be acquired in the assembled route.

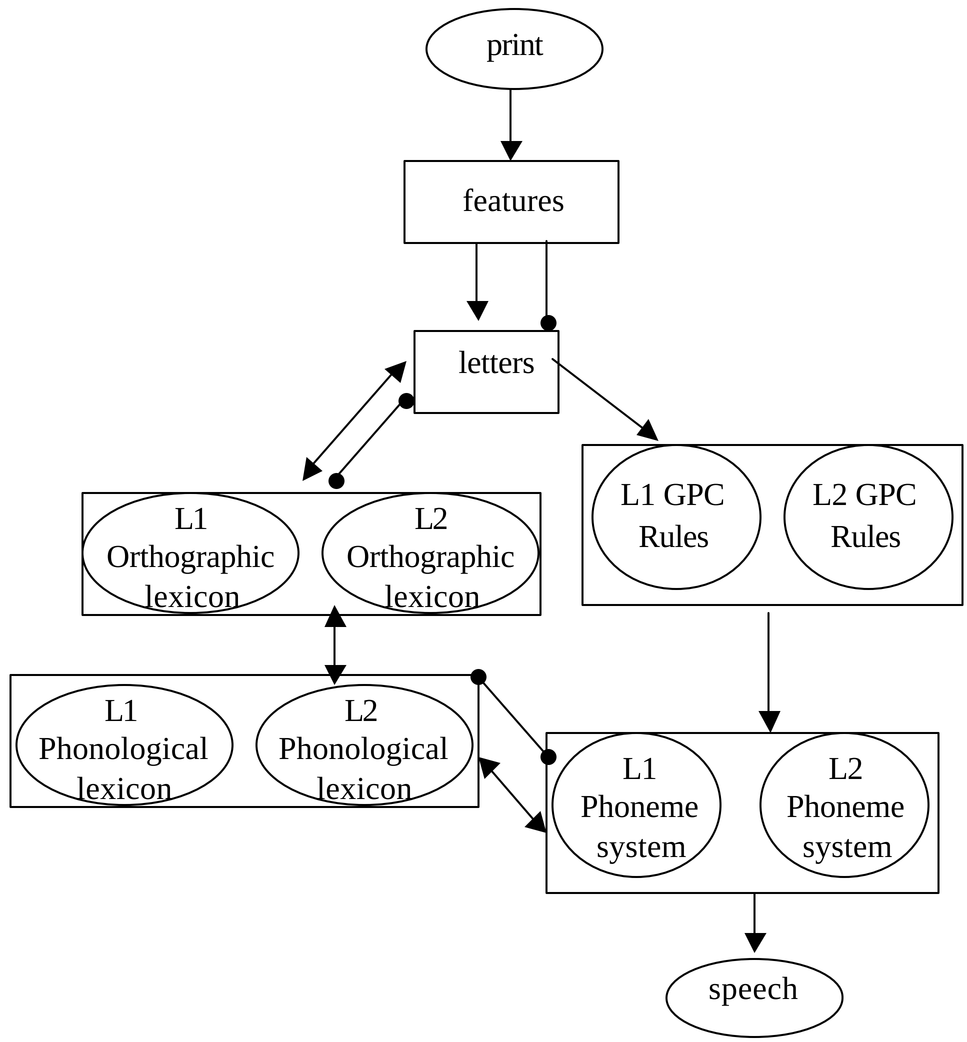


Figure 2.3. A bilingual DRC Model with L2 representations in the direct and assembled routes.

Given the last version of the model, the process of acquiring a second language within a dual-route framework should also look different depending on the native language and the second language of the bilingual. Although there are a number of possible language combinations, four types of language combinations will be addressed: a shallow L1 acquiring a shallow L2, a deep L1 acquiring a shallow L2, a shallow L1 acquiring a deep L2, and a deep L1 acquiring a deep L2.

For languages with shallow orthographies, the dual-route framework suggests that speakers of these languages are more reliant on the assembled route because the mapping from spelling to sound for the language is so consistent. This is not to say that speakers of languages with shallow orthographies never use the direct route to go from spelling to sound because effects attributed to the direct route such as word frequency effects have been found in languages with shallow orthographies such as Spanish (Carreiras, Alvarez, & de Vega, 1993). According to this framework, acquiring a second language with a shallow orthography should be relatively easy given that the native language and the acquired language will mostly rely on rules mapping one letter to one sound. In other words, for the assembled route, both languages will rely on the same grain size of processing. For this case, since the grain size of processing should be similar, at the level of the word, the most difficult aspects to acquire will be cases where there are conflicting spelling-to-sound mappings between the first and the second languages.

In the case of a speaker of a deep orthography acquiring a language with a shallow orthography, the acquisition may be slightly more difficult. In this process the speaker not only has to learn the spelling-to-sound mappings of the new language but also has to start processing words at a smaller grain size. Since even most languages with deep orthographies have some consistency, the learner whose native language has a deep orthography should be able to make this transition somewhat smoothly since there are smaller grain size processing strategies already in place.

Perhaps the more difficult cases according to the dual-route framework are those of native speakers of languages with shallow orthographies learning languages with deep orthographies, and speakers of languages with deep orthographies learning another orthographically deep language. In both instances the learners must acquire the new spelling-to-sound mappings of the new language. For those learners whose native language is shallow, in addition they must learn to process words at a larger grain size and look for consistency in the new language in larger sublexical units. In the case of learners whose first language is deep learning a second language with a deep orthography, these learners already have some processing in place that operates at a larger grain size, however this grain size may not correspond with that of the new language. This point brings us to the two main languages of comparison in this set of experiments: English and German. Although both of these languages have deep orthographies, the processing preferences of both languages are somewhat different. One difference is that although some German letters can map onto more than one sound, overall the language is more consistent than English because the addition of another letter can often resolve the ambiguity of how a given letter is pronounced. However, with English, consistency generally occurs with much larger units. These differences were captured in Zeigler et al.'s (2000) extension of the DRC to German. This extension found a number of differences in the types of rules between English and German (see Table 2.1).

Table 2.1

Rule differences between English and German (Zeigler et al, 2000)

	Language	
Rules	English	German
Single-letter	39	45
Multi-letter	146	48
Context-sensitive	14	38

The rules in Table 2.1 that were extracted from the English and German versions of the DRC demonstrate that spelling-to-sound mappings in German are more consistent than those in English. Not only are there fewer rules in German but the number of single-letter and context-sensitive rules are greater in German than in English. The different proportions of single-letter, multi-letter, and context-sensitive rules across German and English should also have differential consequences for native German speakers acquiring English and native English speakers acquiring German. For proficient German-English bilinguals to process English words like native speakers, they must learn to use larger sublexical units during word recognition and production. Conversely for proficient English-German bilinguals, this large unit processing must be replaced by smaller unit processing in order for processing to look like that of native German speakers.

Regardless of which of the above versions of the Dual-Route model is most likely to represent bilingual language processing, placing language-specific rules in the assembled route makes the clear assumption that language-specific processing strategies are located within the lexicon. It is on this assumption where the Dual-Route model and BIA+ clearly differ.

The Bilingual Interactive Activation + (BIA+) Model

The Identification System of BIA+ closely resembles the direct route of the bilingual Dual-Route models proposed earlier. Both systems contain information on lexical orthography and phonology that is activated in parallel. However the similarities between the two models end with this system. In BIA+, word information processed in the identification system is transferred to the Task Schema level. This level contains a number of higher level processes including decision criteria for making responses, information for how to make task-specific responses (i.e. button presses for lexical decision), and although Dijkstra and van Heuven (2002) do not make this explicit within the model, this level would be the most likely location for language-specific processing information. The main consequence for having language-specific processing information in the Task Schema level is that it makes the assumption that this information influences word recognition and production relatively late in the stream of processing.

General Method and Design

To examine parsing preferences across first and second language speakers of English, three experiments were conducted. Each experiment examined parsing in native speakers of English and proficient bilinguals. The parsing preferences of L2 speakers were then compared with both native English speakers and when possible a group of speakers performing a similar parsing task in the bilinguals' L1. Two tasks from cross-linguistic studies (Goswami et al., 2001; Ziegler et al., 2001), and a third task adapted from a study examining monolingual parsing (Treiman & Chafetz 1987) were used to make these comparisons. The participants, tasks, and the logic for examining the participant groups and tasks will be discussed in the following sections and chapters.

Participants. The participant groups that are common to the three tasks in these experiments were German-English bilinguals and native English speaking controls. This bilingual group was chosen because as mentioned earlier German, like English, is an orthographically deep language but is overall more consistent. This greater consistency has consequences for the grain size of processing for native speakers of these languages (Ziegler et al., 2000; 2001) and may also have consequences on the ease of mapping orthography to phonology in adults (see Goswami et al., 2001 for evidence in children). In addition, the cross-linguistic evidence for differences in parsing preferences across these two languages provides a strong foundation for a bilingual extension of this work. For the first experiment, parsing in Japanese-English bilinguals will also be compared with the German-English bilinguals and the native English controls. This group was added to explore differences in parsing between L2 speakers whose first languages differ in script. However, unfortunately due to the small number of Japanese-English bilinguals at Penn State University, this participant group was only included in Experiment 1. Additional details about the participants for each task will be provided in subsequent chapters.

Experiment 1: Lexical Decision with Pseudohomophones

The first experiment explores parsing through an extension of Goswami et al. (2001). Participants in this experiment spoke English as a first or second language. The critical stimuli were pseudohomophones, nonwords that can be pronounced like real words. For example, in English a possible pseudohomophone is kake (cake). These special nonwords are often used in monolingual studies as well as cross-linguistic studies to determine the degree to which phonology is active during word recognition and production. The general finding is that there is an advantage for naming pseudohomophones over orthographically similar controls (i.e., kake is faster to name than gake). But perhaps the more surprising finding is that pseudohomophones are also harder to reject during lexical decision, a task that does not overtly require retrieving the phonology of the word. In addition, Goswami et al. discovered that when performing lexical decision in their native languages, German children demonstrated a larger disadvantage for rejection pseudohomophones than native English speaking children, most likely because of the greater consistency of German compared to English.

In this experiment, the performance of native and nonnative English speakers was compared on an English lexical decision task. There were two non-native speakers groups. One group was comprised of proficient German-English bilinguals and a second group was made up of proficient Japanese-English bilinguals. German-English bilinguals were an interesting comparison group because like English, German is orthographically deep in that letters can map onto multiple sounds. However as mentioned above, unlike English, German is fairly regular such that given an ambiguously pronounced letter, knowing either the preceding or following letter will generally make the pronunciation unambiguous. Japanese-English bilinguals were another interesting comparison because unlike English, Japanese is a mostly logographic language. In addition the Japanese script that uses the Roman alphabet is orthographically shallow. The distinct properties of German and Japanese enabled the examination of the consequences of the first language on using the parsing preferences in English.

As mentioned earlier, all three language groups made lexical decisions to words and nonwords including pseudohomophones and two critical control groups. The first critical control group was a group of nonwords that were both orthographic and phonological controls (O+P+ controls). These nonwords were neighbors of the base word for the matched pseudohomophone. For example if the pseudohomophone was taip (for tape), the O+P+ control would be fape. A second control group contained nonwords that were orthographically and phonologically dissimilar to the base word of the pseudohomophone (O-P- controls). Returning to the example above, an O-P- control for taip could be zoash (please refer to Table 2.2 for other example stimuli). The logic of utilizing these three nonword conditions will be discussed in Chapter 3.

Table 2.2

Example of Stimulus Materials for Experiment 1

Word	Pseudohomophone	O+P+ control	O-P- control
Fake	Faik	Dake	Koog
Girl	Gerl	Rirl	Cilf

Experiment 2: Cluster lexical decision

Because the pseudohomophone measure of parsing may provide only a partial picture of how parsing unfolds, a second method used to examine this process was a cluster technique similar to one used by Treiman and Chafetz (1987). In the Treiman and Chafetz study, words and nonwords were presented with slashes either after the onset or after the first vowel. For example the word "than" would either be presented parsed as "th//an" or "tha//n." Reaction times for these two different parses were then compared. Using this technique, Treiman and Chafetz (1987) found that for native speakers of English, words that are clustered by onset and rime (i.e. th//an) are recognized more quickly than after the first vowel (i.e. tha//n).

One criticism of this study is that for the stimuli used in this experiment, there were cases where the letters following the slashes formed a word (i.e. st//ale). This raises the possibility that the benefit found for onset-rime parsed words and the disadvantage for onset-rime nonwords could have been due to the presence of these "post-slash" words. In the present experiment, instead of clustering words by slashing, participants were first presented with a fixation appeared on the screen and then either the onset or onset plus vowel. The remaining spaces in the word were either left blank or filled with the same number of asterisks as the remaining letters (see Figure 2.4). After a brief duration, the rime or coda (letters following the vowel) appeared to fill in. Participants were instructed to make a decision after the entire word or nonword was displayed.

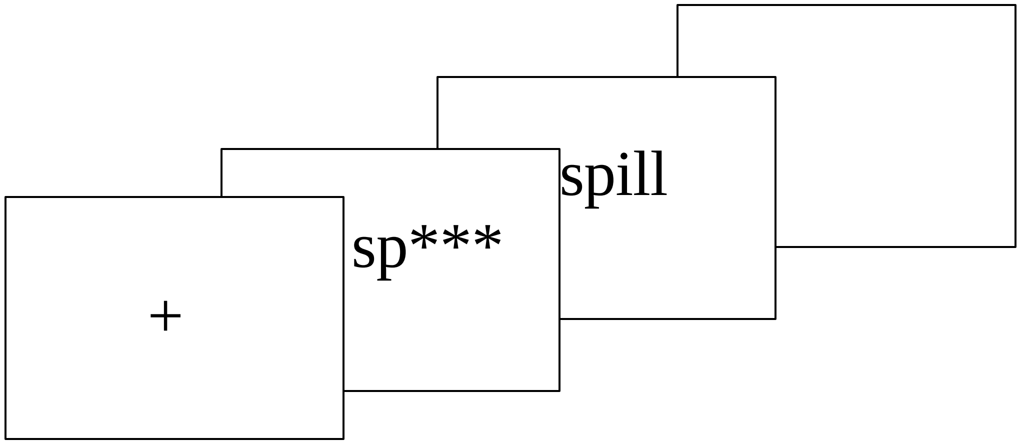


Figure 2.4. Cluster lexical decision task for Experiment 2.

Experiment 3: Cognate Naming

In the third and final experiment, the performance of native English speakers and German-English proficient bilinguals was compared on an English word and nonword naming task. The critical items in this study were cognates across English and German (i.e. hand-Hand). The word and nonword items in this study varied on two critical factors: item length and the number of body neighbors. Words consisted either of 3, 4, 5, or 6 letters and either had many words in English or German that shared the word body, or a small number or words sharing the word body.

In a cross-linguistic study looking at native language, Ziegler et al. (2001) found that word naming performance in native English speakers was more sensitive to the number of body neighbors that an item had, and naming in native German speakers was more affected by word length. The general finding was that words with larger numbers of body neighbors were read more quickly by native English speakers than words with smaller numbers of body neighbors and native German speakers took longer to name six-letter items than three-letter items. The reason given by the authors for these distinct findings across the two languages was that because German is a more consistent language than English, processing in German is able to occur at a smaller grain size than processing in English. In other words, the native German speakers were able to rely on smaller sublexical units when performing the task, thus creating a length effect in German. For native English speakers, processing at a larger grain size allowed naming not to be as affected by word length. However these native speakers did seem to be affected by the frequency of the larger sublexical unit of the rime.

The goal of this experiment was to determine whether proficient German-English bilinguals would reveal native-like lexical parsing strategies in their L2. In other words, this experiment examined whether German-English bilinguals were able to parse the English-German cognates like native speakers of English. In addition, the inclusion of cognates increased the likelihood that German would be active. Past research demonstrating the nonselectivity of lexical access has exploited the presence of cognates across languages to show that cognates are processed differently than noncognate translations. For example, Van Hell and Dijkstra (2003) showed that even when lexical decision is performed in one language only, cognates in other languages in which the participant is proficient are recognized more quickly than noncognate controls. Likewise, the time it takes bilinguals to name cognates even when naming occurs in a blocked language list, is a function of the similarity of the orthography and phonology of the target and nontarget reading of the word (Schwartz et al., 2001). The use of cognates allows the current experiment to also explore whether the acquisition of those strategies constrains the degree of cross-language interactions at the orthographic and phonological levels during bilingual production.

Individual Difference Measures

In addition to the tasks described above participants completed a language history questionnaire (LHQ) and a reading span task based on Waters and Caplan (1996). The LHQ was developed in the Language and Cognition lab at Penn State. For this questionnaire, participants answered a number of questions self assessing their level of language learning and self-rated proficiency in their first and other languages. Although self-assessments are potentially unreliable, the results of past studies suggest that they are useful in matching groups on language experience.

For the reading span task, participants were presented with 2-6 sentences one at a time. The sentences in this task could either be plausible (i.e. “It was the student that took the exam.”), or implausible (i.e. “It was the exam that took the student.”). Participants were instructed to make one response if the sentence was plausible and another response if the sentence was plausible. At the same time, participants were instructed to keep the final word of each sentence in memory. For example if in a block, a participant was presented with the two example sentences given above, he or she would have to remember “exam” and “student” and write these words down in a booklet provided by the experimenter. For the German participants, a German version of the reading span task was created so that they could perform this task in their first language

The reading span task was included both as a measure of cognitive load for participants and as a tool for determining the locus of the parsing effects in the three tasks. The motivation for using this task as a measure of cognitive load is that comprehending and producing in a second language is a cognitively demanding task. Therefore, it is possible that similarities could exist in the processing of low reading span native English speakers (who find processing English cognitive demanding) and the non-native speaker groups.

The reading span task can also be used to help determine the locus of parsing effects across the three experimental tasks. There is evidence that individuals with a large working memory capacity are able to suppress irrelevant information more effectively than individuals with smaller working memory capacity. Because this suppression mechanism operates fairly late in processing, evidence for effects of working memory capacity in the following experiments would suggest that the effects of lexical parsing are relatively late (Michael & Golan, submitted). Failure to find effects of working memory capacity would suggest that any effects of lexical parsing are early, bottom-up processes.

Methods of Analysis and General Predictions for Experiments 1-3

The performance of native speakers of English and nonnative speakers of English on these tasks will be compared mainly through analysis of variance. Within these analyses we will look for differences in parsing strategies between native and nonnative speakers of English and also differences within the nonnative speakers to determine if there are differences according to the script and orthographic depth of the native language.

Table 2.3

Summary of Proposed Experiments and Tasks

	Goal	Tasks	Participants
Exp. 1	To identify and compare lexical parsing strategies for native and nonnative speakers of English through the use of pseudohomophones	-lexical decision	-Native speakers of English -German-English bilinguals -Japanese-English bilinguals
Exp. 2	To identify and compare lexical parsing strategies for native and nonnative speakers of English through the using German-English cognates	-naming	-Native speakers of English -German-English bilinguals
Exp. 3	To compare processing strategies by examining the consequences of violating a natural English parse	-"cluster" lexical decision	-Native speakers of English -German-English bilinguals

The three experiments were selected to examine a number of variables that have been demonstrated to distinguish lexical parsing across language groups. In Experiment 1, the main variables were orthographic and phonological overlap between nonwords and actual words for three groups of nonwords. This experiment examined the consequences of this degree of overlap between native speakers of English, German, and Japanese. These three languages vary in the consistency for spelling-to-sound mapping with English having the least consistent mappings and Japanese having the most consistent mappings. In addition, this experiment explored how differences in L1 script may influence parsing by comparing the performance of Japanese participants, native speakers of a mainly logographic language with the performance of native speakers of English and German.

If native speakers of English use larger units to recognize and produce words, then for the pseudohomophone task, there should be smaller differences in the response times for pseudohomophones and O+P+ than O-P- nonword controls and according to the Dual-Route model, participants should also take longer to identify both of these groups as nonwords compared to O-P- controls.

If both the native German and native Japanese participants utilize this L2 parsing preference during English lexical decision, the pseudohomophone disadvantage should be similar across all three groups. However if these two native speaker groups are relying in L1 parsing preferences the lexical decision task should reveal a larger pseudohomophone disadvantage for both native German and native Japanese speakers due to the greater consistency of their native languages. In other words, if nonnative speakers are using smaller processing units, and are able to efficiently map spelling to sound in English, there should be larger inhibitory effects for pseudohomophones than for the other two types of nonwords for lexical decision. In addition, because Japanese is a more consistent language than German, if these two groups of nonnative speakers are using L1 processing strategies, the pseudohomophone effect should be larger for the native Japanese group, than for the native German group.

In the cluster lexical decision task, parsing preferences are measured by comparing the processing time for strings of letters parsed by onset and rime, to those parsed after the first vowel. In addition this experiment attempts to examine the time course of this processing by manipulating the SOA between the presentation of the first and second parts of these clustered stimuli (see Figure 2.4). If the results from this modified task replicate the findings of Treiman and Chafetz (1987), native speakers of

English should find words separated by onset and rime easier to recognize than words that are separated after the vowel even though in the latter case more information is being presented early on. If non-native speakers are not using the L2 processing preferences of the larger units of the onset and rime when processing English words, then the difference in processing time for onset-rime parsed stimuli compared to nononset rime parsed stimuli should differ in magnitude and or direction from the native English speakers. More specifically L2 English speakers should either demonstrate no difference in response times for the two different types of clustering, or perhaps even an advantage for the onset-vowel cluster.

The third and final experiment examines lexical parsing through a production task. For this third experiment, instead of making lexical decisions, participants actually name words and nonwords aloud. This third task measures the effects of body neighborhood (N) size and also word length on naming. One other manipulation for this experiment was the presence of English-German cognates, a manipulation that was included to maximize the probability that the first language of the native German speaker group would be active. If native speakers of German are relying on L1 processing strategies for this task, then like the German participants in the Ziegler et al. (2001) study who named either identical or nearly identical words in German, native German speakers should show larger effects of word length than native English speakers despite the fact that they are naming words in English and not German. In addition, if these L2 speakers are relying on L1 strategies, they should show smaller effects for body N size than native speakers of English. However, if the pattern of results for the native German speakers does not differ from the native English speakers, then this suggests that these highly proficient non-native speakers were able to not only acquire L2 processing strategies, but also use these strategies in the presence of strong L1 cues.

Chapter 3: Experiment 1: Lexical Decision with pseudohomophones

Because English is an orthographically deep language, recent evidence has demonstrated that native speakers of English parse words into large sublexical units such as the orthographic rime and use this information during word recognition and production (Kessler & Treiman, 1997; Treiman et al., 1995; Ziegler & Perry, 1998). In contrast, speakers of orthographically shallow or more consistent languages seem to rely on smaller grapheme-to-phoneme correspondences. However, what is not clear is whether proficient non-native speakers of English adopt the same lexical parsing preferences as native speakers when recognizing and producing English words. The aim of this first experiment was to examine the lexical processing of non-native speakers to determine whether the parsing strategies they use to process English words differs from native English speakers, and if so to determine whether processing differences depend on the native language.

The method used to examine parsing preferences in Experiment 1 was to compare the magnitude of pseudohomophone effects in English across three native speaker groups using lexical decision. The main responses of interest were response latencies to three types of nonwords: pseudohomophones and two types of nonword controls. One third of the nonwords were pseudohomophones, nonwords that are pronounced like actual words (see Table 2.2). For example, the nonword “kake” shares its pronunciation with the English word “cake.” Previous studies using English word naming have found that pseudohomophones such as “kake” are named more quickly than nonpseudohomophonic nonwords like “dake” (Seidenberg, Peterson, MacDonald, & Plaut, 1996; Taft & Russell, 1992). In contrast to the findings in word naming, in lexical decision, pseudohomophones are generally slower to reject and have higher error rates than nonpseudohomophonic nonwords especially when the pseudohomophone is an orthographic neighbor of the word with which it shares a pronunciation (Rastle & Coltheart, 1994). For example native English speakers take longer to reject pseudohomophones like “kake” versus nonwords like “gake.”

Cross-linguistically, studies have demonstrated that pseudohomophones are not processed in the same way across all languages. As mentioned in Chapter 1, the developmental study by Goswami et al. (2001) showed that like adults, English children showed an advantage for naming pseudohomophones over nonpseudohomophonic nonwords, whereas German children did not show this advantage. These effects occurred even though the children in both groups were not yet skilled readers.

The difference in advantage for this special type of nonword has been attributed to the differences in consistency between English and German (Goswami et al., 2001). Because German is more consistent than English, there is evidence that processing is based on a smaller grain size than English (Ziegler et al., 2001). In other words, the consistency of German allows words and nonwords to be processed via smaller sublexical units, and processing at these smaller units makes the phonological computation for pseudohomophones and nonpseudohomophonic nonwords more similar. This ease of computation through smaller units negates any advantage for having the lexical phonology of the baseword of the pseudohomophone (e.g., cake for kaik). Therefore, there is no difference in naming latencies or errors for these two types of nonwords. However, because English is less consistent, phonology is computed at a larger grain-size and is less efficient. Because this mapping process is less efficient, the lexical phonology of the baseword does affect processing and produces an advantage for pseudohomophones over other nonwords for naming.

In lexical decision, greater efficiency in computing phonology from orthography appeared to create a processing disadvantage for rejecting pseudohomophones for German-speaking children. Because they were able to compute the word-like phonology of the pseudohomophones, it also made it more difficult to make a “no” response, resulting in greater error rates for pseudohomophones compared to the English-speaking children. Like the cross-linguistic differences in naming, the difference in processing time for these two types of nonwords is thought to be related to the greater consistency of German than English (Goswami et al., 2001).

The current experiment is an extension of the Goswami et al. (2001) study in two respects. The first goal was to determine if the pattern of the pseudohomophone effect found for native English speaking children could be replicated in adults for both lexical decision and naming. The second, and more important goal, was to determine if the pseudohomophone effect found for native English speakers was similar in proficient non-native speakers of English and, if not, whether it depends on the degree of proficiency in English and/or lexical preferences associated with the non-native speaker’s L1.

To examine this question, three groups of English speakers were examined. The first was a group of native English speakers who were not proficient in an L2, and therefore functionally monolingual. The second group was comprised of German-English bilinguals. This group was chosen because like English, German is an orthographically deep language, however unlike English, German is more consistent. In addition, as mentioned earlier a number of cross-linguistic studies have found differences in processing preferences between English and German (Goswami et al., 2001; Ziegler et al., 2001). The third group was made up of Japanese-English bilinguals to determine whether script differences between the L1 and L2 affect how second language parsing preferences are acquired.

In addition to pseudohomophones, two critical nonword control conditions were included. In one condition nonwords were orthographically and phonologically similar (O+P+) to the base word of the pseudohomophone (see Table 2.2). For example, if the critical pseudohomophone was *faik*, with a base word of *fake*, then the O+P+ control would be *dake.* The other nonword condition contained words that were orthographically and phonologically dissimilar (O-P-) to the base words of the pseudohomophones and to other real words. Given the above example, an O-P- control for *faik* could be the nonword *koog*.

The characteristics of these three nonword conditions provide a basis on which comparisons can be made both within and across language groups. The O+P+ controls were the most “word-like” nonwords of the three nonword conditions. These nonwords differed from the base word by only the first letter, and were therefore nearly identical orthographically and phonologically to the base word. During processing, these nonwords should activate the base word as well as its orthographic neighbors providing strong evidence that these items could be real words.

Nonwords in the pseudohomophone condition were orthographically dissimilar to the base word, but phonologically identical. These nonwords should eventually activate the phonological representation of the base word but not the orthographic representation.

Therefore these nonwords should produce less overall activation during processing than the O+P+ controls.

Finally the O-P- controls were the least word like of the three control groups. These words were dissimilar from the base word and other words in both spelling in sound. Of the three nonword conditions, this nonword control was least likely to activate the baseword representation.

Having both the O+P+ nonword condition in addition to the O-P- nonword condition was important because in most experiments that examine the processing of pseudohomophones, performance on pseudohomophones is generally compared only to nonwords that are orthographically and phonologically dissimilar to most real words (the O-P- condition). Including nonwords that are orthographically and phonologically similar to the base word will allow for a more sensitive analysis of the pseudohomophone effect to determine whether having an identical phonetic match to an item in the lexicon has greater consequences than having both a highly similar phonetic and orthographic match.

Method

Participants

Three groups of participants were recruited for this study: native speakers of English (n=41), late German-English bilinguals (n=43), and late Japanese-English bilinguals (n=17). The group of native English speakers and Japanese-English bilinguals were recruited from The Pennsylvania State University. The proficient German-English bilinguals were recruited from the University of Muenster and the Max Planck Institute of Cognitive Neuroscience in Leipzig. In all cases, bilinguals were identified as individuals who had acquired their L2 past early childhood and who use L2 at a high level of proficiency. Native English speakers were recruited through the psychology department subject pool and received course credit for their participation. Bilinguals were recruited through e-mail and printed advertisements and were paid for their participation.

Materials

The materials used in this experiment were those used by Goswami et al. (2001). These items were chosen so that a direct comparison could be made between the adult data from the current experiment and the child data from Goswami et al. (2001). In addition, the word items from this study and the base words from which the nonword conditions were created were all words that should be familiar to non-native speakers of English. Because the native English speaker group in the Goswami et al. study spoke British English, some pseudohomophone items needed to be changed to make the pronunciation conform to American English (see Appendix A). Stimulus materials were blocked by the number of syllables. Each syllable block contained 24 words and 24 nonwords. For the 24 nonwords in each block, 8 were O-P-, 8 were O+P+, and 8 were pseudohomophones. The order of the blocks was counterbalanced across participants and within each block items were presented in a random order.

Procedure

Participants were tested one at a time while seated at the computer. On each trial, participants were presented with a fixation in the center of the screen for 500 ms followed by a letter string forming either a word or a nonword. Participants were instructed to press a key marked "2" if the string of letters formed an English word and a key marked "1" for nonword responses. The "2" key was the rightmost key on a 5 key button box and the "1" key was the leftmost key on the button box. The unused keys in the center of the button box were covered to make only the response buttons visible. In addition, only responses made by these two buttons were recorded by the experiment program. Participants were instructed to respond as quickly as possible but not at the expense of accuracy.

After the completion of the lexical decision task, native English speakers and native German speakers completed a reading span task based on Waters and Caplan (1996) in their L1. For the native German speakers, a German version of the task was created.[i] In this task, participants were seated at a computer and presented with 20 blocks of 2-6 plausible and implausible sentences. Participants were allowed up to 5 seconds to make the plausibility judgment and after 2-6 sentences, participants were then asked to recall the final word from each sentence from the block. Responses were recorded in a booklet provided by the experimenter.

At the end of the experimental session, participants completed a short language history questionnaire in which they were asked questions about their proficiency in their first language and other languages that they had acquired.

Results and discussion

For this experiment, each language group was first analyzed individually to examine the pattern of effects of nonword type and syllable length for that group. In the final analyses, the effects of syllable length and nonword type will be examined including all three language groups to systematically examine these effects across the three language groups. For the primary analyses, each language group was separated into two groups using an individual difference measure. To determine whether working memory capacity influences parsing, the native English speaker group was divided into two groups using the number recalled on the reading span task, and the two non-native groups were divided into high and low proficiency groups. Memory span data was also collected from the native German speaker group, however for the primary analysis, to determine whether level of proficiency affects parsing, and in order to maintain consistency with the native Japanese speaker group, the German group was only divided by proficiency. The effect of memory span on performance for the native German speaker group will be discussed at the end of the chapter.

Participant exclusion and language history questionnaire data

Prior to the analysis, the participants' language history questionnaires were examined to make certain that all participants were native speakers of their language group. Four participants from the native English speaker group were removed from the analysis because another language was spoken in the home in addition to English. In the native German speaker group, one participant was removed because another language was spoken in the home, a second participant was removed because her first language was Romanian, and a third was removed because she was born in England to German parents. In the native Japanese speaker group, one participant was excluded because she was an early Japanese-English bilingual.

Table 3.1

Mean Language History Questionnaire data for Native German and Japanese Speaker Groups

Native Language group	Proficiency Rating	Age (yrs.)	Reading (out of 10)	Writing (out of 10)	Speaking (out of 10)	Comprehension (out of 10)	Months of Immersion
German	Low	26.5	6.3	5.1	6.1	6.6	6.9
		(6.9)	(1.0)	(0.8)	(0.9)	(1.3)	(1.6)
	High	27.1	8.1	7.8	8.0	8.6	8.0
		(6.5)	(0.9)	(0.9)	(0.8)	(0.9)	(1.7)
	Mean	26.9	7.4	6.8	7.3	7.8	7.3
Japanese	Low	28.9	6.6	6.6	5.3	5.4	30.4
		(2.5)	(1.5)	(1.6)	(1.3)	(1.1)	(25.0)
	High	27.4	8.1	7.7	7.9	8.1	58.8
		(6.6)	(1.3)	(1.9)	(0.7)	(0.4)	(22.5)
	Mean	28.2	7.4	7.1	6.6	6.8	44.6

Standard deviations in parentheses

In addition to the language requirement, participants were excluded from the analysis if they did not achieve at least 75% accuracy for words and nonwords and 80% overall accuracy. Five participants from the native German speaker group failed to meet these criteria as well as 2 participants in the native Japanese speaker group. Also, one additional participant from the German-English bilingual group who met both the language and accuracy criteria was excluded because during the experiment she would often look to the person who accompanied her to the experiment for help on the task. Excluding these participants left 37 native English speakers, 34 German-English bilinguals, and 14 Japanese-English bilinguals to be included in the following analyses.

For the remaining participants, the mean ratings in English given on the language history questionnaire over reading, writing, speaking, and comprehension for the German-English bilinguals was 7.3 (out of 10) and for Japanese-English bilinguals, the average rating was 7.0. This rating difference was not statistically significant ($t_{46} = 0.89$, $p < 0.4$).

Reading span results

To determine whether or not working memory has an effect on the pattern of results native English and native German participants completed a working memory task in their L1. Japanese participants did not complete this task because currently there is no version of this task in their L1. Participants had a maximum of five seconds to make a plausibility judgment for each sentence and then at the end of each block, they were asked to recall the last word of each sentence and record those words in a booklet provided by the experimenter. Outliers for this task were identified by finding the reaction time value 2.5 standard deviations above and below the mean for both plausible judgments and implausible judgments for each participant. The number of words recalled for each participant was calculated by tallying the number of words recalled only for trials that were correct and also not outliers. Intrusions were words that participants recorded in the booklet that were not actual target words. The mean values for the plausibility judgments, number of words recalled, and intrusions can be found in Table 3.2.

For the purposes of the main analyses the variable of interest was number of words recalled. Participants in the native English speaker group were also split into two groups according to the number of words recalled on the reading span task. The total number of words possible for recall on this task is 80, and the average number of words recalled for this group was 43 (*SD* = 8.3). Participants were placed in the low span group if they recalled 0-39 words and in the high span group if they recalled 40 or more words.

Table 3.2

Reading Span results for native English and native German speakers

Language Group	Mean Plausible RT (ms)	Mean Implausible RT (ms)	Mean correct judgments (out of 80)	Mean Words recalled (out of 80)	Mean number of Intrusions
English	3101	3250	68	43	6
	(348.3)	(312.8)	(7.9)	(8.3)	(4.5)
German	2956	3052	65	50	3
	(300.0)	(296.5)	(6.0)	(8.8)	(2.3)

Standard deviations in parentheses

For the follow-up analysis for the native German participants, the dividing mark between the low span and the high span group was raised to 47 words recalled because on average this group recalled seven more words than the native English group. The results of the German participants suggest that these participants either had an overall higher working memory capacity due to their bilingualism, or that the German version of the task was perhaps slightly easier than the English version. To create high and low memory span groups for the native German speakers, those participants that recalled 0-47 words were included in the low span group and those recalling 48-80 words were included in the high span group.

Calculation of outliers

Outliers for the lexical decision task for each participant were identified by taking the mean response time over words and nonwords and computing 2.5 standard deviations above and below the mean. Items outside of this range were trimmed and not included in the analyses. These trimmed trials accounted for less than 3% of the data.

Native English Speaker Group and Comparison to Goswami et al. (2001) Child Data

Reaction Time for nonwords

Reaction time data for the native English speaking adults were analyzed using a 3 (syllable length) X 3 (nonword type) X 2 (memory span) ANOVA (see Table 3.3 for a listing of the means). This analysis revealed a main effect for syllable length by

Table 3.3

Mean RT (ms) and accuracy (out of 100%) by syllable length and nonword type for Experiment 1

	RT			Accuracy		
Syllable length	O-P-	O+P+	PSH	O-P-	O+P+	PSH
1 syllable	645	728	641	98.5	93.6	98.9
	(23.9)	(40.7)	(24.6)	(0.11)	(0.19)	(0.11)
2 syllable	673	714	707	98.1	97.0	98.8
	(36.0)	(29.5)	(31.5)	(0.15)	(0.16)	(0.08)
3 syllable	710	722	728	97.8	95.4	82.4
	(30.4)	(23.2)	(24.5)	(0.12)	(0.14)	(0.15)
Nonword type mean (RT)	676	721	692	98.2	95.3	93.3

Standard errors in parentheses

participants, (*F1*(2,70) = 3.46, *MSE* = 17264.7, $p < .04$), and by items (*F2*(2,63) = 7.17, *MSE* = 2326.8, $p < .003$), and a main effect for nonword type by participants, (*F1*(2,70) = 13.58, *MSE* = 3920.8, $p < .001$), and by items (*F2*(2,63) = 5.38, *MSE* = 2326.8, $p < .009$). The syllable length X nonword type interaction was significant for participants (*F1*(4,140) = 4.35, *MSE* = 5396.3, $p < .003$) but marginal by items (*F1*(4,63) = 2.20, *MSE* = 2326.8, $p < .08$). In the participant analysis, there was no difference between the two memory span groups, ((*F1*(1,35) = 0.005, *MSE* = 207788, $p < .95$), and memory span did not interact with either syllable length (*F1*(2,70) = 0.28, *MSE* = 17264.7, $p < .76$), or nonword type (*F1*(2,140) = 0.61, *MSE* = 3920.8, $p < .54$). Because there was no effect for memory span in the participant analysis, this variable was dropped from further analyses.

Post hoc analyses revealed that the main effect for syllable length for participants was due to an average of 49 ms slower response times for 3 syllable words than for 1 syllable words. After Bonferroni correction, this difference was marginally significant ($t_{36} = 2.34$, $p < .08$). For nonword type, on average O-P- nonwords were responded to 48 ms faster than O+P+ nonwords ($t_{36} = 5.31$, $p < .002$), O-P- nonwords were marginally faster than PSHs by 17 ms ($t_{36} = 2.46$, $p < .06$), and response times for PSHs were on average 31 ms faster than O+P+ nonwords (p < .004). These main effects were also qualified by a significant interaction of syllable length and type. O+P+ nonwords were significantly slower by 87 ms than O-P- nonwords ($t_{36} = 3.90$, $p < .003$), and 90 ms than PSHs ($t_{36} = 3.87$, $p < .002$), only for 1 syllable nonwords. These differences were not significant in the other two syllable conditions.

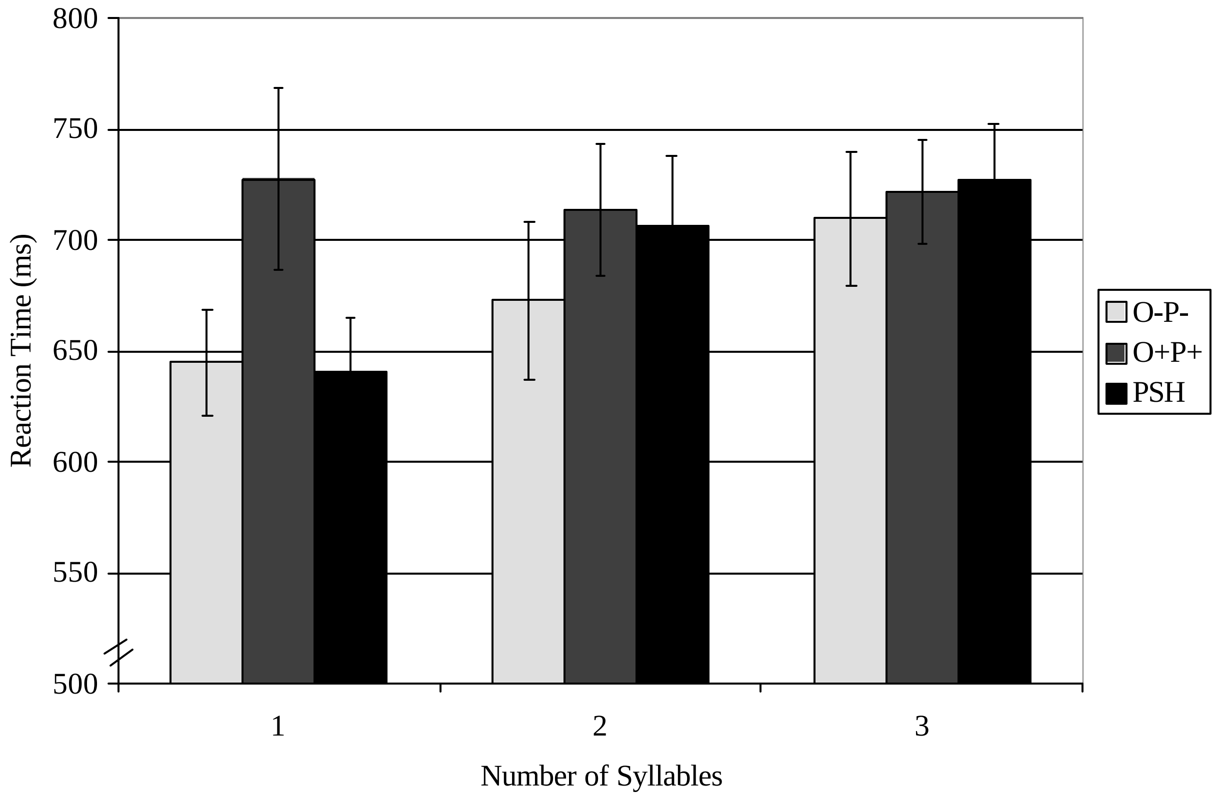


Figure 3.1. Mean reaction times (ms) for Native English speaker group by number of syllables and nonword type for Experiment 1.

Accuracy for nonwords

Accuracy for nonwords on this task was high (*M* = .96, *SD* = .08). However despite this high level of accuracy, there were still some differences within the conditions. Accuracy was analyzed using a 3 X 3 X 2 (syllable length X nonword type X memory span) ANOVA (see Table 3.3 for accuracy by condition). Due to the high levels of accuracy on this task, prior to the analysis, the data were recoded using the arcsin transform (2*arcsin $\sqrt{x}$). This analysis revealed a main effect for syllable for participants ($F1(2,70) = 7.81$, $MSE = 0.23$, $p < .002$), and items ($F2(2,63) = 4.51$, $MSE = 0.08$, $p < .02$), a main effect for nonword type for participants ($F1(2,70) = 5.86$, $MSE = 0.15$, $p <$

.005) or but not items (*F2*(2,63) = 2.21, *MSE* = 0.08, $p < .2$), and an interaction for syllable and type for participants (*F1*(4,140) = 10.38, *MSE* = 0.15, $p < .001$), and items (*F2*(4,63) = 4.51, *MSE* = 0.08, $p < .003$). As in the reaction time analysis, there were no differences in overall accuracy between the two memory span groups (*F1*(1,35) = 0.62, *MSE* = 0.28, $p < .5$), and span failed to interact with either syllable length or type.

The main effect for syllable length arose from a 5% difference in accuracy between the one-syllable and three-syllable length conditions (97% vs. 92%) ($t_{36} = 2.96$, $p < .02$), and a 6% difference between the two- and three-syllable length conditions (98% vs. 92%) ($t_{36} = 3.25$, $p < .01$). For nonword type, O-P- nonwords had significantly higher accuracy (98%) than both O+P+ nonwords (95%) ($t_{36} = 2.59$, $p < .05$), and PSHs (93%) ($t_{36} = 3.29$, $p < .01$). However these main effects were qualified by a significant interaction. Inspection of this interaction found that for 1-syllable words, O-P- nonwords were marginally more accurate than O+P+ nonwords ($t_{36} = 2.86$, $p < .07$), and PSHs were significantly more accurate than O+P+ items ($t_{36} = 3.50$, $p < .01$), (99% and 99% vs. 94% respectively). For 2-syllable words there were no significant differences between word types. However for 3 syllable words, there was significantly lower accuracy for PSH nonwords (82%) than for O-P-(98%) ($t_{36} = 6.47$, $p < .002$), and O+P+ (95%) nonwords ($t_{36} = 4.31$, $p < .002$).

Overall, these results suggest that for shorter words, the word-like phonology from the pseudohomophones did not interfere as much as the word-like orthography of the O+P+ nonwords. However for multi-syllabic nonwords where processing was extended for a longer period of time, the word-like phonology of the pseudohomophones interfered more as the words became longer.

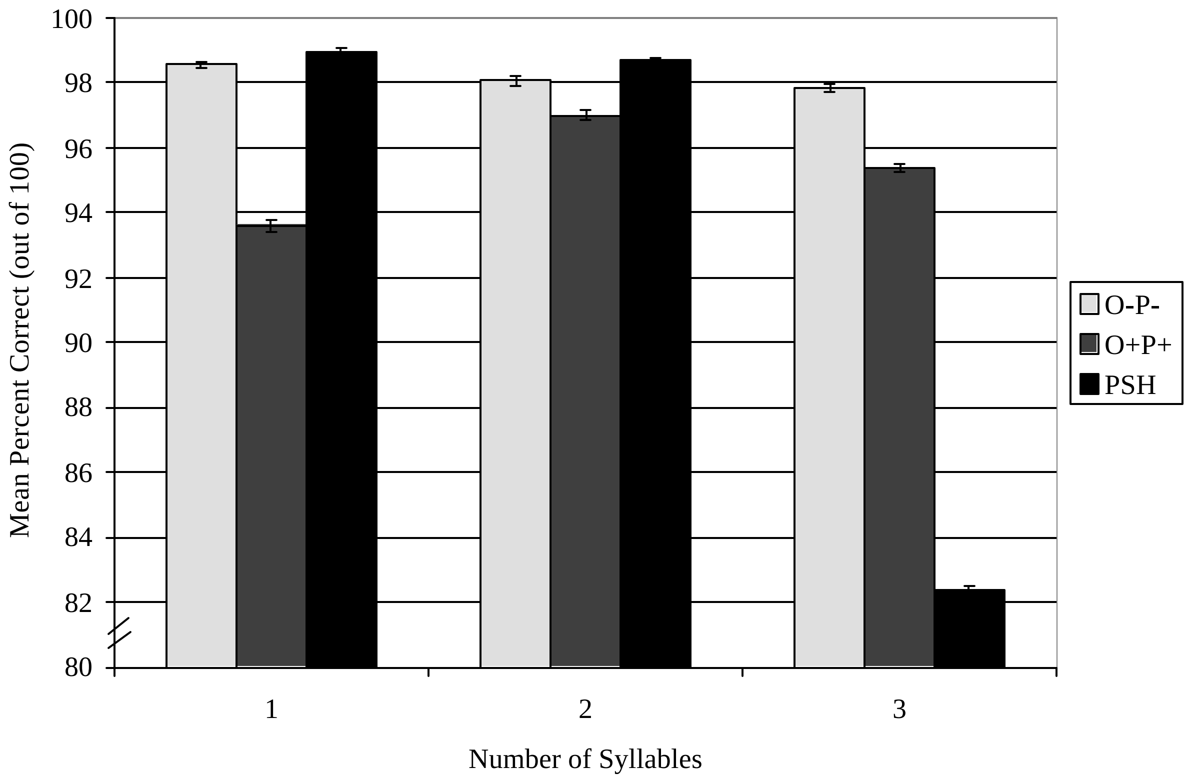


Figure 3.2. Mean accuracy for Native English speaker group by syllable length and nonword type for Experiment 1.

Reaction time for words

Although the word data were less critical for this study, they were analyzed to ensure that word responses were faster than nonword responses and to also examine the effect of syllable length across the three groups of participants. The mean RTs for the nine nonword groups (3 syllable length X 3 nonword type) and 3 word groups were averaged to create a grand word mean and nonword mean for each participant. These means were then compared by a paired-samples t-test. This analysis found that nonword responses were on average 117 ms slower than word responses ($t_{36} = 6.26$, $p < .001$).

The word responses were then analyzed using a 3 (syllable length) X 2 (memory span) ANOVA. This analysis found a main effect for syllable length by participants (*F1*(2,70) = 35.96, *MSE* = 1748.9, $p < .001$), and by items (*F2*(2,69) = 24.70, *MSE* = 2040.5, $p < .003$). Post hoc analyses revealed that this main effect arose from significant differences between all three syllable groups. On average, 1 syllable words were 34 ms faster than 2 syllable words ($t_{36} = 3.42$, $p < .006$), and 90 ms faster than three syllable words ($t_{36} = 8.28$, $p < .001$), and 2 syllable words were 56 ms faster than 3 syllable words ($t_{36} = 6.52$, $p < .001$).

There was no overall effect of working memory span in the participant analysis (*F1*(1,35) = 0.07, *MSE* = 22915, $p < .8$). There was also no interaction between working memory span and syllable length (*F1*(2,70) = 2.06, *MSE* = 1748.9, $p < .2$). Therefore this variable was again dropped from the item analysis.

Table 3.4

Mean RT (ms) for native English speaker word data in Experiment 1

Syllable length	1	2	3
RT	539 (12.6)	573 (15.6)	629 (17.3)

Standard errors in parentheses

Accuracy for words

The mean accuracy for words and nonwords was first compared. The mean accuracy value for the nonword conditions was computed by averaging the accuracy percentages for the nine nonword conditions (3 syllable length X 3 nonword type). For the mean word accuracy, the accuracy percentages for the 1, 2, and 3 syllable words were averaged. These grand percentage means were then compared using a paired-samples t-test. This analysis found that there were significant differences between the word and nonword items for the native English speaker group ($t_{36} = 2.81$, $p < .009$). This difference was due to a 2% accuracy advantage for words over nonwords.

The word data were then transformed through the arcsin and analyzed using a 3 (syllable length) X 2 (memory span) ANOVA. This analysis revealed a main effect for syllable length by participants ($F1(2,70) = 6.53$, $MSE = .06$, $p < .003$), but only marginal by items ($F2(2,69) = 2.43$, $MSE = .06$, $p < .1$). Post hoc analyses showed found no significant differences between the syllable conditions however this main effect seems to be due to greater similarity in accuracy between 1 and 2 syllable words than 1 and 3 syllable words and 2 and 3 syllable words.

There were no overall differences between the two memory span groups ($F1(1,35) = 0.18$, $MSE = 0.08$, $p < .7$). The interaction between syllable length and memory span was also not significant ($F1(2,35) = 2.22$, $MSE = .06$, $p < .2$). Therefore this variable was not included in the item analysis.

Table 3.5

Mean accuracy for word data (out of 100%) in Experiment 1

Syllable length	1	2	3
Mean Percent Correct	98.6 (.06)	98.9 (.06)	95.8 (.04)

Standard errors in parentheses

Overall, this analysis shows that like the nonwords, the words in this experiment showed effects of syllable length for both accuracy and reaction time. Perhaps the more surprising result from the word analysis is that all three syllable groups were significantly different from each other unlike the nonword analysis where significant differences existed only between the 1 and 3 syllable conditions. If phonological information for nonwords can only be computed serially, then the expected finding would be a larger syllable effect for nonwords than words. However these findings suggest that either the information for these nonwords were not being computed serially, or the nature of the lexical decision task did not require the phonology to be fully retrieved. These points will be discussed in greater depth in the general discussion.

Comparison to Goswami et al. (2001)

In the study conducted by Goswami et al. (2001), children ages 7-9 who were in the early stages of reading made lexical decisions to the same words and nonwords used in Experiment 1. Although the higher level of variability in the child data makes it difficult to compare the two groups directly, there are a number of similarities. For reaction time in the child data, there were no significant differences in reaction time for any of the three nonword conditions. However across the three syllable conditions, the children performing this task seemed to respond most quickly to the O-P- condition and somewhat slower to the O+P+ nonwords and PSHs (see Table 3.6 for condition means). Although the pattern between the children and the adults is not exactly the same there are some similarities between the two groups. For both groups, O-P- nonwords have the fastest response times followed by the PSHs and then the O+P+ nonwords. This suggests that the word-like orthography and phonology of the O+P+ nonwords were distracting for both the unskilled child readers and the more skilled adult readers. However the most striking difference between the adult and child data is the size of the pseudohomophone effect. Although this effect was not significant in Goswami et al.'s analysis, there is a 137 ms difference between the O-P- condition and the PSH condition. This difference was much smaller for adults (16 ms) and not significant. The dissimilarity in the magnitude of the differences between these two nonword conditions suggests that the children in the Goswami et al. study may have been more dependent on the phonological representations of the nonwords when performing this task. The smaller difference between O-P- nonwords and PSHs for adults suggests that the phonological information had less of an effect on response time. Given that the children were most likely poorer readers than the adults in Experiment 1 and that research has demonstrated that poorer readers are often influenced more by phonology (Jared, Levy, & Rayner, 1999), this explanation seems likely.

Table 3.6

Nonword RT (ms) and accuracy data (out of 100%) from Experiment 1 and Goswami et al. (2001) collapsed across syllable length

	Experiment 1		Goswami et al. (2001)	
	RT	Accuracy	RT	Accuracy
O-P-	676	98.2	1697	91.7
O+P+	721	95.3	1891	82.7
PSH	692	93.3	1834	80.0

Although there was some overlap in the reaction time data for the child and adult groups, for accuracy, the similarities between the two groups are more striking. Although Goswami did not find any significant differences between conditions because of high variability, overall, the children had the highest levels of accuracy for the O-P- nonwords and similar levels of accuracy for the O+P+ nonwords and PSHs.

This pattern was closely mirrored by the adults in Experiment 1. Like the children, the adults were most accurate for the O-P- nonwords, and least accurate for the O+P+ nonwords and PSHs.

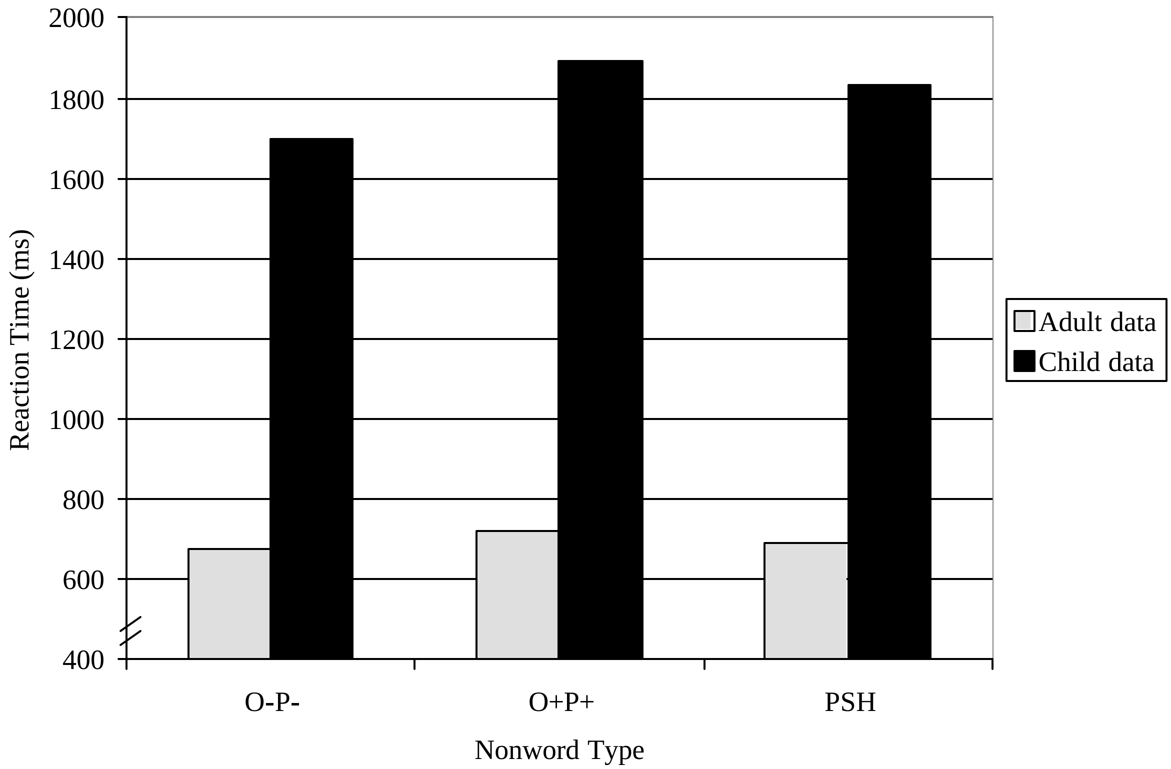


Figure 3.3. Mean RT data for nonword type collapsed over syllables for adults in Experiment 1 and native English speaking children (Goswami et al., 2001).

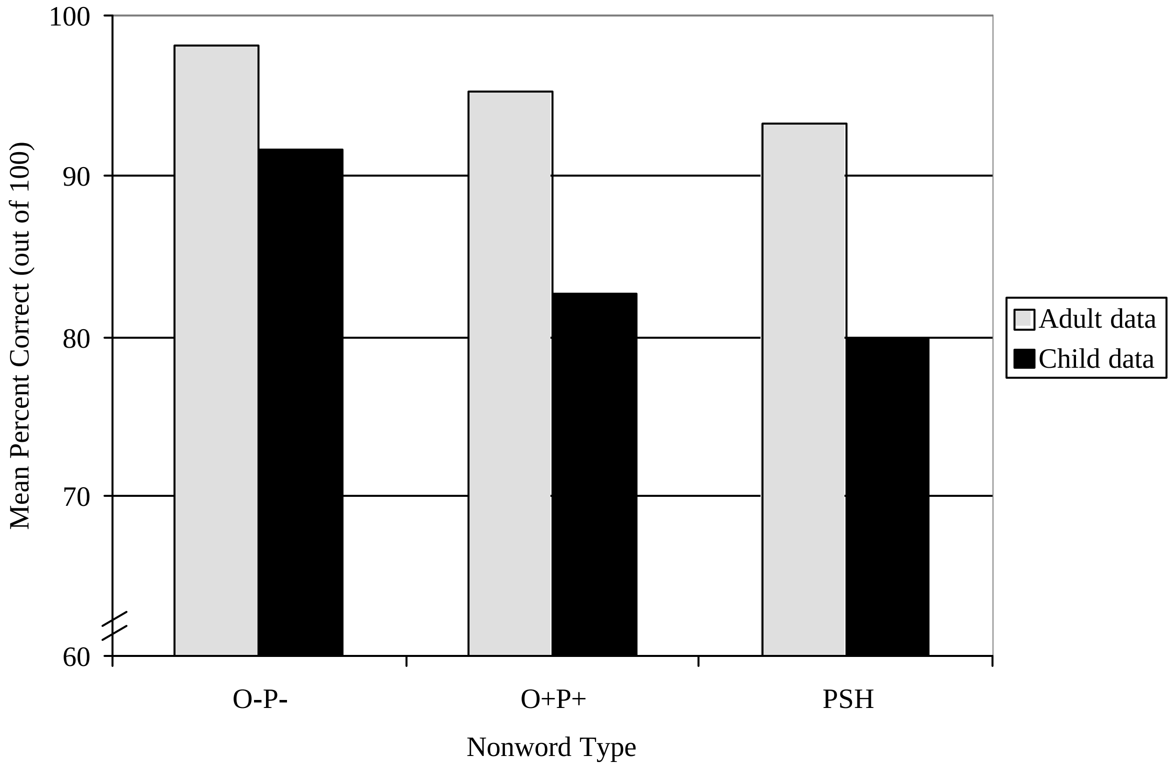


Figure 3.4. Mean Accuracy collapsed over syllables for adults in Experiment 1 and native English speaking children by nonword type (Goswami et al., 2001).

Together with the reaction time data, these results suggest that although the children may have been more reliant on phonological information during this task, phonology was active during the decision process for both groups and disrupted processing. However from both the child and the adult data, having the exact phonology of the PSHs did not appear to be significantly more disruptive than having the similar orthography and phonology of the baseword in the case of the O+P+ nonwords.

Native German Speaker Group

Reaction Time data for nonwords

The analysis for the native German speaker group was similar to the analysis for the native English speakers except in this case, since span was not a critical variable in the English group, and this participant group had relatively high memory span, the individual difference measure of interest was proficiency. The results for the memory span data for this participant group will be discussed later in this chapter. Reaction time data for the native German speaker group was analyzed by a 3 (syllable length) X 3 (nonword type) X 2 (proficiency rating) ANOVA. Proficiency ratings were based on participants' responses to questions about proficiency in English reading, writing, speaking, and comprehension. Ratings were made on a scale from 1-10 and the ratings in the four categories mentioned above were averaged to gain a mean proficiency rating.

Table 3.7

Mean RT (ms) and accuracy (out of 100%) by syllable length and nonword type for Native German speakers for Experiment 1

Syllable length	O-P-	O+P+	PSH	O-P-	O+P+	PSH
1 syllable	787	848	772	97.4	93.3	95.0
	(38.8)	(49.1)	(36.8)	(.14)	(.22)	(.15)
2 syllables	773	829	837	98.9	96.2	96.3
	(40.8)	(39.7)	(45.8)	(.09)	(.17)	(.14)
3 syllables	835	858	863	98.9	97.6	90.7
	(50.2)	(52.1)	(50.3)	(.11)	(.16)	(.19)
Nonword type mean	798	845	824	98.4	95.7	94.0

Standard errors in parentheses

Participants whose mean proficiency rating was below 7.0 were categorized as low proficiency, and participants whose average rating was 7.0 and above were categorized as high proficiency.

Unlike the native English speaker group, the native German speakers did not show a syllable length effect for participants (*F1*(2,64) = 1.89, *MSE* = 30270, $p < .2$), or for items (*F2*(2,63) = 0.7, *MSE* = 6105.3, $p < .6$). This result is perhaps surprising considering that length has been shown to affect processing in German (Ziegler et al., 2001). There was however a main effect for nonword type for participants (*F1*(2,64) = 13.02, *MSE* = 3940.6, $p < .001$), and items (*F2*(2,63) = 3.97, *MSE* = 6105.3, $p < .03$). This main effect was due to a 47 ms advantage for processing O-P- nonwords over O+P+ nonwords ($t_{33} = 6.15$, $p < .002$) and a 25 ms advantage for processing O-P- nonwords over PSHs ($t_{33} = 2.82$, $p < .03$). This main effect is qualified by an interaction of syllable and type for participants (*F1*(4,128) = 4.37, *MSE* = 5565.7, $p < .03$), but not items (*F2*(4,63) = 1.44, *MSE* = 6105.3, $p < .3$). Examination of the interaction found that for 1 syllable words, there were no significant differences between processing time for O-P- nonwords and PSHs ($t_{33} = 0.84$, $p < 1.0$), however there were significant processing advantages for O-P- nonwords ($t_{33} = 3.43$, $p < .02$) and PSHs ($t_{33} = 3.64$, $p < .006$) over O+P+ nonwords (62 and 77 ms respectively). For two syllable words, O-P- nonwords were processed significantly faster than both O+P+ nonwords ($t_{33} = 4.88$, $p < .002$) and PSHs ($t_{36} = 4.09$, $p < .002$) (56 and 65 ms respectively). For 3 syllable words, there were no significant differences between the three nonword categories. In addition, there were no differences between proficiency for groups in overall performance in the analysis by participants (*F1*(1,32) = .19, *MSE* = 519143, $p < .7$), and proficiency failed to interact with either syllable length and nonword type (*F1*(4,128) = .37, *MSE* = 5565.7, $p < .9$). Because there were no effects of proficiency in the subject analysis, this variable was dropped in further analyses.

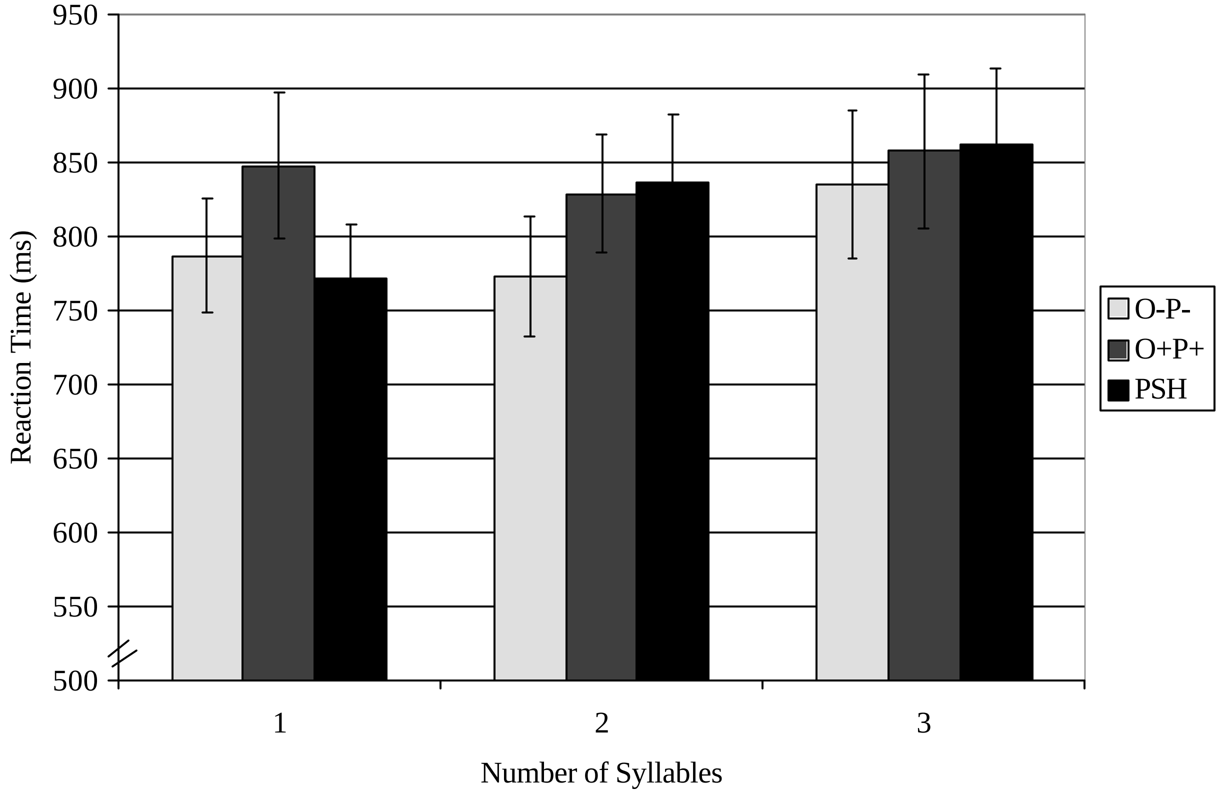


Figure 3.5. Mean RT for Native German speakers by syllable length and nonword type.

Accuracy for nonwords

Accuracy for the native German speaker group was also high (*M* = .96, *SD* = .02). The 3 (syllable length) X 3 (nonword type) X 2 (proficiency) ANOVA using arcsin transformed accuracy scores once again found no main effect for syllable length for participants (*F1*(2,64) = 1.01, *MSE* = 0.25, $p < .4$), or items (*F2*(2,63) = 1.14, *MSE* = .09, *p* < .4). There was a main effect for nonword type for participants (*F1*(2,64) = 6.68, *MSE* = 0.21, *p* < .003), and items (*F2*(2,63) = 4.40, *MSE* = .09, *p* < .02). Post Hoc analyses revealed that O-P- nonwords were responded to with higher accuracy than both O+P+ nonwords (t_{33} = 2.54, *p* < .05) and PSHs (t_{33} = 3.98, *p* < .002) (98% vs. 94% respectively). However the difference between O+P+ nonwords and PSHs was not significant (t_{33} = 1.03, *p* < 1.0). In addition, there was a significant interaction between syllable length and nonword type for participants (*F1*(4,128) = 3.07, *MSE* = .12, *p* < .02), but not items (*F2*(4,63) = 1.32, *MSE* = .09, *p* < .3). Examination of this interaction found that there were only differences between nonword types for 3 syllable words. For this syllable group accuracy for PSHs was lower than O-P- nonwords (t_{33} = 4.56, *p* < .003)

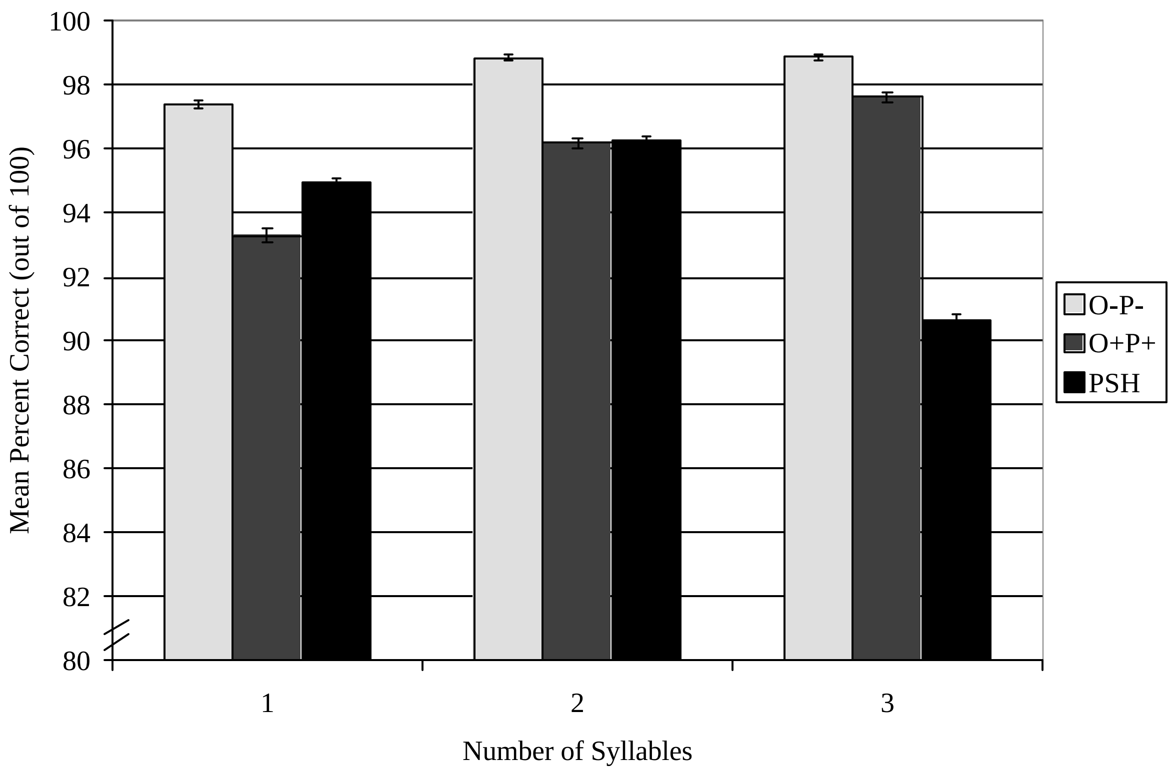


Figure 3.6. Mean Accuracy for Native German speakers by syllable length and nonword type.

and O+P+ nonwords (t_{33} = 2.98, $p < .02$) (91% vs. 99% and 98% respectively). Once again, there were no differences in the participant analysis for performance between the two proficiency groups ($F(1,32) = 0.18$, $MSE = .36$, $p < .7$), and proficiency did not interact with syllable length ($F1(1,32) = 0.09$, $MSE = 0.25$, $p < .1$), or nonword type ($F1(1,32) = 0.27$, $MSE = 0.21$, $p < .8$). Therefore proficiency was not included in the item analysis.

Although the pattern for the native German participants is not exactly like the native English speakers, the patterns for the two groups are remarkably similar. One surprising finding was that the native German speakers did not show a syllable length effect for reaction time or accuracy even though past studies have shown that length affects processing in German. This finding suggests that this speaker group is processing English in a different way than they are processing German. In addition to the lack of a clear syllable effect there are many other processing similarities between the native German and native English group. Both groups demonstrated overall differences for nonword type with O-P- nonwords showing significant processing advantages over O-P- nonwords and PSHs (this second difference was only marginal for the native English group). The interaction between syllable length and nonword type also looked very similar across the two groups. Both showed processing advantages for O-P- nonwords over both O+P+ nonwords (marginal in English native speakers), and PSHs for 1 syllable words. For longer words, there were significant processing disadvantages for PSHs over O-P- nonwords and O+P+ nonwords. One interesting difference between the two groups however was for the native English group, this disadvantage was manifested for 3-syllable words and for the native German group, this disadvantage was significant for 2-

syllable words. This difference suggests that although the processing strategies of the native German group look very similar to those of the native English group, the efficiency of recoding spelling to sound may be allowing the PSH effect to appear earlier for the German group than for the English group.

Reaction time for words

Word responses for the native German speaker group were compared with nonword responses using a paired samples t-test. As expected, this analysis found that word responses were faster than nonword responses ($t_{33} = 5.89$, $p < .001$) by an average of 152 ms.

The word responses were then analyzed using a 3 (syllable length) X 2 (proficiency rating) ANOVA. This analysis found a main effect for syllable length by participants (*F1*(2,64) = 11.00, *MSE* = 2815.44, $p < .001$), but not items (*F2*(2,69) = 1.57, *MSE* = 12044.5, $p < .3$). Post hoc analyses found that the main effect by participants was due to significantly faster responses for 1 syllable words than 3 syllable words ($t_{33} = 4.52$, $p < .003$) (64 ms) and significantly faster responses for 2 syllable words than 3 syllable words ($t_{33} = 3.18$, $p < .01$) (44 ms).

There were no overall differences for participants between the higher and lower proficiency groups (*F1*(1,32) = 0.71, *MSE* = 45721, $p < .5$), and proficiency rating did not interact with syllable length (*F1*(1,32) = 1.17, *MSE* = 3351.15, $p < .3$). Therefore proficiency was not included in the item analysis.

Table 3.8

Mean RT (ms) for native German speaker word data in Experiment 1

Syllable length	1	2	3
RT	642	658	702
	(18.9)	(22.0)	(27.5)

Standard errors in parentheses

Accuracy for words

The mean accuracy for words and nonwords was compared using a paired-samples t-test. This analysis found a marginal difference in accuracy between the words and nonwords for this experiment ($t_{33} = 1.99$, $p < .06$). This marginal difference was due to a 2% average higher accuracy for words over nonwords.

Accuracy rates were then examined using a 3 (syllable length) by 2 (proficiency rating) ANOVA. This analysis showed a main effect for syllable by participants ($F1(2,64) = 28.62$, $MSE = .03$, $p < .001$), but not items ($F2(2,69) = 0.33$, $MSE = .12$, $p < .8$). Examination of the main effect for participants revealed that for the native German speakers, 3 syllable words were significantly more accurate than both 1 syllable ($t_{33} = 5.18$, $p < .001$) and 2 syllable words ($t_{33} = 6.59$, $p < .001$).

There were no overall accuracy differences between the higher and lower proficiency groups, ($F(1,32) = 2.18$, $MSE = .05$, $p < .2$), and proficiency did not interact with syllable length ($F(2,64) = 1.04$, $MSE = .03$, $p < .4$). Once again this variable was dropped from the item analysis.

Table 3.9

Mean accuracy (out of 100%) for native German speaker word data in Experiment 1

Syllable length	1	2	3
Mean Percent Correct	93	92	98
	(.02)	(.02)	(.06)

Standard errors in parentheses

The reaction time findings for the word data for native German speakers are also similar to the native speakers of English. Although there were no differences between the 1 syllable and 2 syllable conditions, the native German speakers were significantly slower to respond to 3 syllable words compared to 1 and 2 syllable words. For accuracy the groups did differ. The native German group made significantly more errors for 3 syllable words than 1 syllable words and 2 syllable words. Although there were some overall differences for the native English group, no significant differences appeared in the post hoc analyses. The differences present at these syllable lengths suggests either that some of the native German speakers were less familiar with the two and three syllable words or that their efficiency in recoding spelling to sound for the nonwords made them more hesitant and less accurate in labeling these items as words.

Native Japanese Speaker Group

Reaction Time for nonword data

Like the native German speaker group, the reaction time data for the native Japanese speaker group were analyzed by a 3 (syllable length) X 3 (nonword type) X 2 (proficiency rating) ANOVA. This analysis revealed a main effect for syllable length by participants ($F1(2,24) = 3.85$, $MSE = 103753$, $p < .04$), and by items ($F2(2,63) = 7.71$, *MSE* = 91487.1, $p < .002$). This main effect arose from marginally slower response times for three syllable words than one syllable words by an average of 194 ms ($t_{13} = 2.70$, $p <$ .06). There was no main effect for nonword type by participants (*F1*(2,24) = 0.61, *MSE* = 52594, $p < .6$), or items (*F2*(2,63) = 0.80, *MSE* = 91487.1, $p < .5$). There was no interaction between syllable length and nonword type by participants (*F1*(4,48) = 0.36, *MSE* = 27970, $p < .9$), or items (*F2*(4,63) = 0.16, *MSE* = 91487.1, $p < 1.0$).

There was no overall difference between the high proficiency group and the low proficiency group for subjects (*F1*(1,12) = 1.71, *MSE* = 3522999, $p < .22$) but there was a difference between the two groups by items (*F2*(1,63) = 98.56, *MSE* = 74661.8, $p <$ .001). The item analysis found surprisingly that on average low proficiency participants were 452 ms faster than high proficiency participants in rejecting nonwords ($t_{71} = 9.42$, $p < .001$). Proficiency also interacted with syllable length in the item analysis (*F2*(2,63) = 4.97, *MSE* = 74661.8, $p < .01$) but not the participant analysis (*F1*(2,24) = 2.26, *MSE* = 103753.7, $p < .2$); and both syllable length and nonword type (*F1*(4,48) = 2.59, *MSE* = 27970, $p < .05$) in the participant analysis, but not in the item analysis (*F2*(4,63) = 1.35, *MSE* = 74661.8, $p < .3$). Inspection of the proficiency by syllable length interaction found that the overall syllable effect was only present in the high proficiency group and not the low proficiency group. For the high proficiency group, 3 syllable words were responded to slower than 1 syllable words by an average of 415 ms ($t_{46} = 4.40$, $p < .004$). This same conclusion was reached in examining the syllable length by word type by proficiency interaction because no differences across word types reached significance after Bonferroni correction.

Table 3.10

Mean RT (ms) by syllable length, nonword type, and proficiency for Native Japanese speakers for Experiment 1

Proficiency rating	Syllable length	O-P-	O+P+	PSH
High	1 syllable	1238	1277	1310
		(206.4)	(234.6)	(222.5)
	2 syllables	1367	1527	1352
		(295.0)	(302.5)	(243.8)
	3 syllables	1652	1643	1556
		(266.3)	(225.9)	(239.4)
	Nonword type mean (RT)	1419	1482	1406
Low	1 syllable	913	1079	931
		(206.4)	(233.6)	(222.5)
	2 syllable	980	933	1088
		(295.0)	(302.5)	(243.8)
	3 syllable	967	980	1115
		(266.3)	(225.9)	(239.4)
	Nonword type mean (RT)	953	997	1045

Standard errors in parentheses

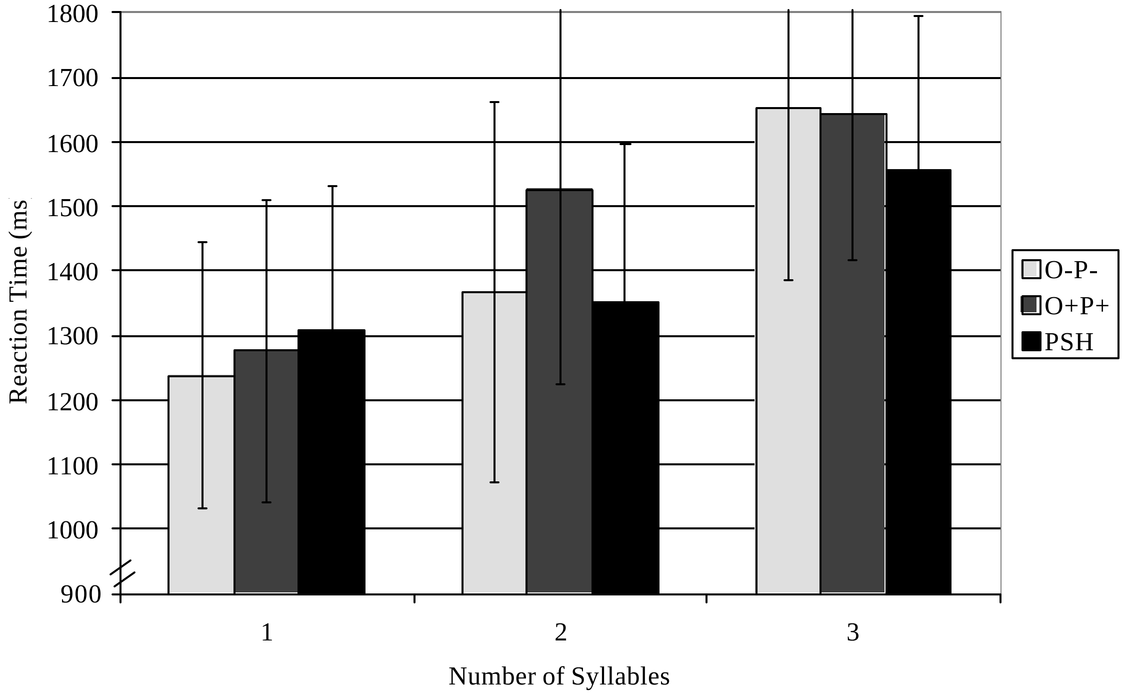


Figure 3.7. Mean RT for high proficiency Native Japanese speakers by syllable length and nonword type.

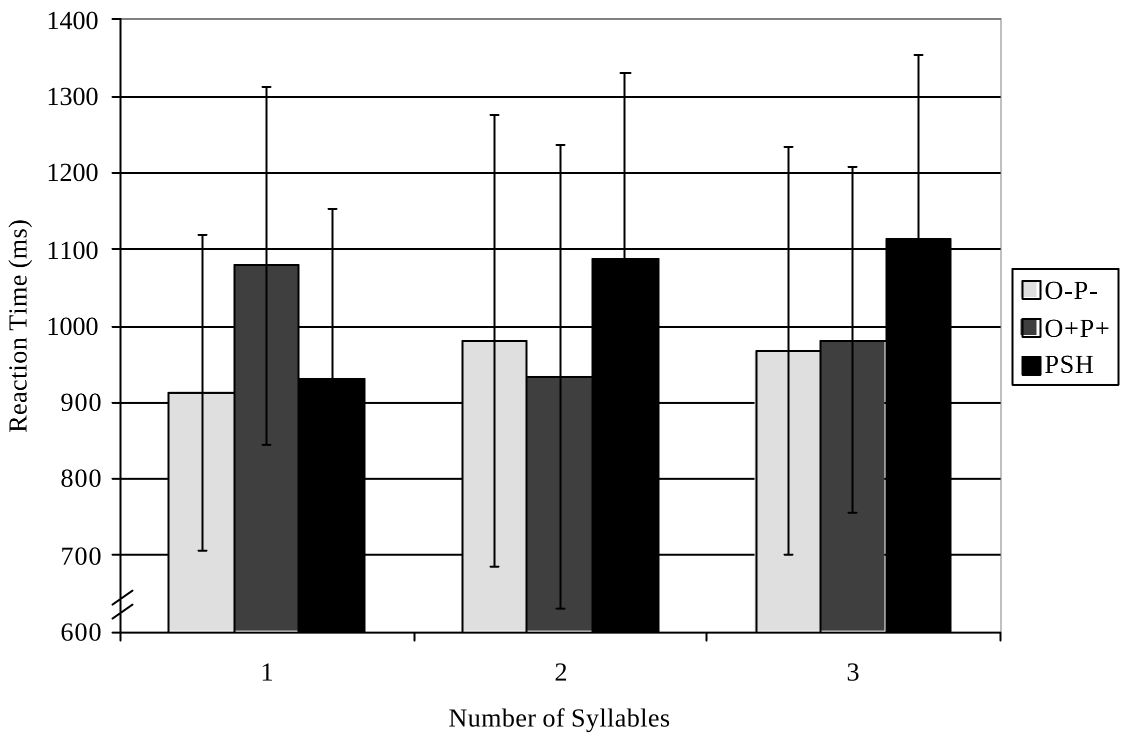


Figure 3.8. Mean RT for low proficiency Native Japanese speakers by syllable length and nonword type.

Accuracy for nonwords

Like the native English and German speaker group, accuracy for the native Japanese speaker group was high (*M* = .95, *SD* = .04). Even with this high level of accuracy there were some differences across the critical variables and groups. A 3 (syllable length) X 3 (nonword type) X 2 (proficiency rating) ANOVA examining the arcsin transformed accuracy data found no main effect for syllable length by participants ($F1(2,24) = 0.40$, $MSE = 0.27$, $p < .7$), or by items ($F2(2,63) = 0.22$, $MSE = 0.18$, $p < .9$). There was a main effect for nonword type by participants ($F1(2,24) = 6.76$, $MSE = 0.16$, $p < .005$), however this effect was not significant by items ($F2(2,63) = 2.22$, $MSE = 0.18$, $p < .2$). This main effect was the result of higher response accuracy to O-P- items (98%) than PSH items (91%). Overall there were no differences in accuracy between O+P+ nonwords and the O-P- and PSH nonwords. There was no interaction of syllable and nonword type for participants ($F1(4,48) = 1.98$, $MSE = 0.15$, $p < .2$), or items ($F2(4,63) = 1.07$, $MSE = .18$, $p < .5$), and there was no overall effect of proficiency ($F(1,12) = 1.26$, $MSE = 0.40$, $p < .3$), and proficiency failed to interact with syllable length ($F(2,24) = 0.46$, $MSE = 0.27$, $p < .7$), and nonword type ($F(4,48) = 1.36$, $MSE = 0.16$, $p < .3$). Finally there was no three-way interaction of syllable length, nonword type, and proficiency rating ($F(4,48) = 0.62$, $MSE = .15$, $p < .7$). Therefore this variable was not included in the item analysis.

Table 3.11

Mean accuracy (out of 100%) by syllable length and nonword type for Native Japanese speakers for Experiment 1

Syllable length	O-P-	O+P+	PSH
1 syllable	99	92	96
	(.27)	(.47)	(.32)
2 syllables	99	96	96
	(.09)	(.17)	(.14)
3 syllables	99	98	91
	(.11)	(.16)	(.19)
Nonword type mean (RT)	98	96	94

Standard errors in parentheses

The nonword data from this native speaker group is very strikingly different from both the native English speaker group and the native German speaker group. For the native Japanese speaker group, syllable length played a much larger role for processing nonwords than for nonword processing in the native German speaker group. This difference may arise from the high consistency of all three Japanese alphabets, the logographic alphabets, hiragana and katakana, and romaji, the Roman alphabetic script for Japanese. This script has a much more consistent spelling to sound mapping than German, therefore this greater consistency may have caused the syllable length effect for this group.

Perhaps the most surprising result from the nonword data for this group is that the low proficiency group was faster and also appeared to be more accurate than the high proficiency group. Although post hoc analyses did not detect a difference between the O+P+ and PSH nonwords, the slower RTs and lower accuracy for these nonwords suggest that the high proficiency group may have been using a different strategy than the low proficiency group. Because this group most likely has greater knowledge of English than the low proficiency group, they took extra time rejecting the very word-like O+P+ nonwords and at times mistook them for real English words. In addition, this native speaker group seemed to be less homogeneous than the other two speaker groups as evidenced by the high degree of variability in the data. Taking into consideration the low number of participants, it is difficult to draw conclusions about the strategies that this group may have been using when processing these nonwords. However for both the low and high proficiency groups, there do seem to be some speed differences between the three syllable conditions which are most likely arising from bringing L1 processing strategies to this L2 lexical decision task.

Reaction time for word data

Overall, the word data and its comparison to the nonword data seem to support these claims. The average word and nonword responses were compared for native Japanese speakers using a paired-samples t-test. This analysis found that like native English and German speakers, native Japanese speakers were significantly faster when making word decisions than nonword decisions ($t_{13} = 3.06$, $p < .01$). What is striking about this comparison is that the difference between word and nonword responses

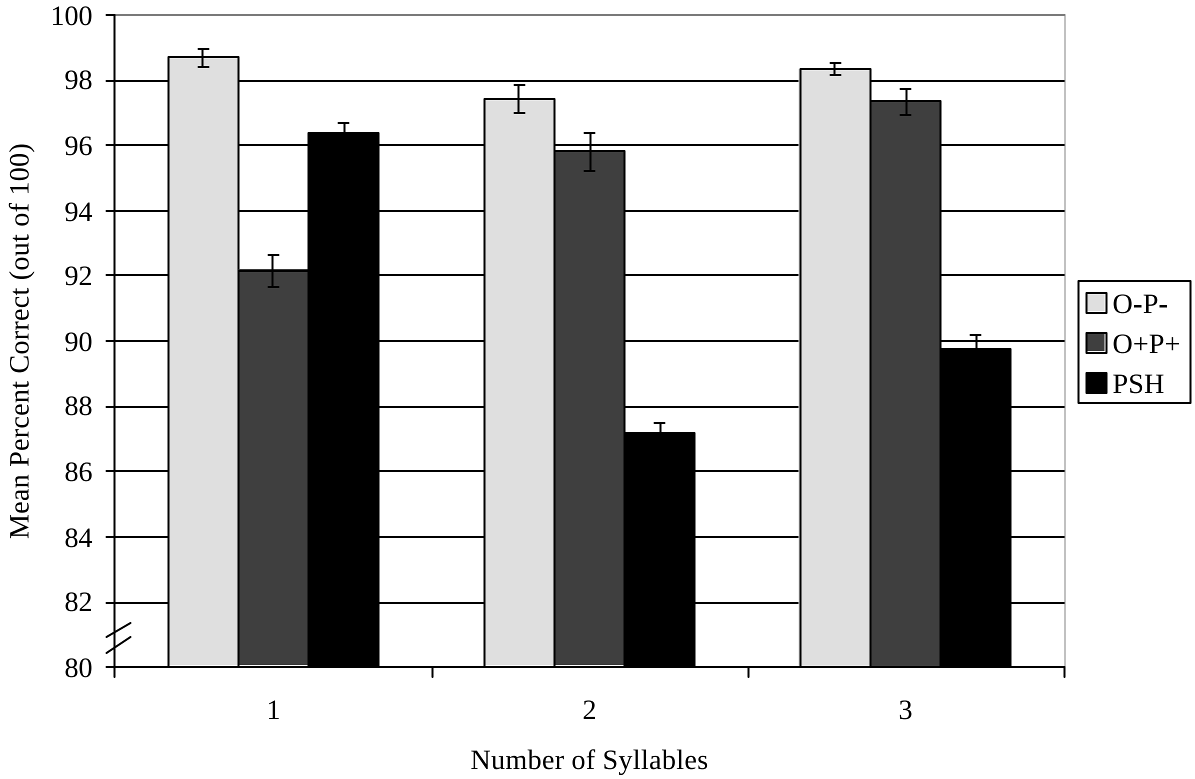


Figure 3.9. Mean Accuracy for Native Japanese speakers by syllable length and nonword type.

for this native speaker group was 344 ms, which is more than twice the difference for the native English and German groups. However considering the nonword results from this native speaker group and the hesitancy of the high proficiency group for making nonword decisions for O+P+ and PSH nonword conditions, this large difference is not surprising.

The word data was then analyzed using a 3 (syllable length) X 2 (proficiency group) ANOVA. This analysis found that for native Japanese speakers there was a main effect for syllable length by participants ($F1(2,24) = 9.47$, $MSE = 32925$, $p < .001$), and items ($F2(2,69) = 14.38$, $MSE = 39787.5$, $p < .001$). This main effect arose from significant differences between 1 and 2 syllable words, and 1 and 3 syllable words. Post hoc analyses revealed that on average, response times to 1 syllable words were 201 ms faster than 2 syllable words, and 292 ms faster than 3 syllable words.

There was no overall effect of proficiency (F(1,12) = 1.70, MSE = 202357, $p < .3$) and proficiency did not interact with syllable length (F(2,24) = 1.29, MSE = 32925, $p <$ .3).

Table 3.12

Mean RT (ms) for native Japanese speaker word data in Experiment 1

Syllable length	1	2	3
RT	708 (53.4)	909 (78.0)	1000 (101.1)

Standard errors in parentheses

Accuracy for words

For the native Japanese speaker group, the difference in accuracy between words and nonwords was marginal (t_{13} = 1.83, $p < .09$). This marginal difference was the result of a 4% greater accuracy for words over nonwords. A 3 (syllable length) X 2 (proficiency group) ANOVA on the arcsin transform was used to examine the accuracy rates for the word items. This analysis revealed a main effect of syllable length for participants ($F1$(2,24) = 3.69, MSE = 0.07, $p < .05$), but not items ($F2$(2,69) = 1.93, MSE = .18, $p <$ .2). This main effect was due to greater accuracy for the 1 syllable condition than the 2 syllable condition. For the Japanese participants, there was a main effect for proficiency ($F1$(1,12) = 11.89, MSE = 0.08, $p < .005$) by participants and items ($F2$(1,69) = 7.71, MSE = .15, $p < .008$). On average, participants in the higher proficiency group had a 4% higher accuracy rate for words than the lower proficiency group. For the native Japanese speaker group, proficiency also marginally interacted with syllable length by participants ($F1(2,24) = 2.90$, $MSE = 0.07$, $p < .08$), and items ($F2(2,69) = 3.08$, $MSE = .15$, $p < .06$). This marginal interaction was the result of accuracy differences for 1 and 2 syllable words only for the lower proficiency groups.

Table 3.13

Mean accuracy (out of 100%) for native Japanese speaker word data by syllable and proficiency rating in Experiment 1

Syllable length	1	2	3
Accuracy: lower proficiency	99 (0.39)	88 (.16)	95 (.20)
Accuracy: higher proficiency	100 (0.17)	100 (.14)	98 (.54)

Standard errors in parentheses

Overall the word data for this group seems to suggest that the native Japanese speaker group was using smaller units of processing than the other two native speaker groups. Although syllable length effects were found in the word data for native English speakers and also native German speakers, these differences were much smaller than those in the native Japanese speaker group. In addition, for the native Japanese group, the main effect of proficiency in the accuracy analysis with higher proficiency participants being more accurate, does support the claim that these participants were more cautious about rejecting nonwords which made reaction times slower and accuracy lower for these participants than the lower proficiency participants.

Analysis across all language groups

The analyses of the individual language groups have shown that there are a number of differences between the three groups. However, overall the native English and native German speaker groups seem more similar in both reaction time and how they process English words and nonwords than the native Japanese speaker group. The following analyses look at these three speaker groups together to gain a sense of how similar these language groups are in processing syllables and the three nonword types.

Analysis across language groups: Syllable length

Reaction time for nonwords. To determine how these three groups differ in processing 1, 2, and 3 syllable words an nonwords, a 3 (syllable length) X 3 (language group) analysis was conducted using the mean response time for each syllable condition

Table 3.14

Mean RT (ms) for syllable type by language group

Language group	1 syllable	2 syllable	3 syllable
English	674	696	720
	(47.6)	(57.2)	(55.9)
German	803	813	855
	(49.7)	(59.7)	(58.4)
Japanese	1124	1208	1319
	(77.4)	(93.1)	(90.9)

Standard errors in parentheses

(collapsed across nonword type) as the dependent measure. This analysis revealed a main effect of syllable by participants ($F1(2,164) = 13.59$, $MSE = 12498.0$, $p < .001$), and items (*F2*(2,69) = 8.36, *MSE* = 24547.4, *p* < .001). A main effect of language group by participants (*F1*(2,82) = 13.98, *MSE* = 295822, *p* < .001), and items (*F2*(2,138) = 402.62, *MSE* = 12731.2, *p* < .001), and a significant interaction by syllable length and language group by participants (*F1*(4,164) = 2.54, *MSE* = 12497.0, *p* < .05), and items (*F2*(4,138) = 5.07, *MSE* = 12731.2, *p* < .001).

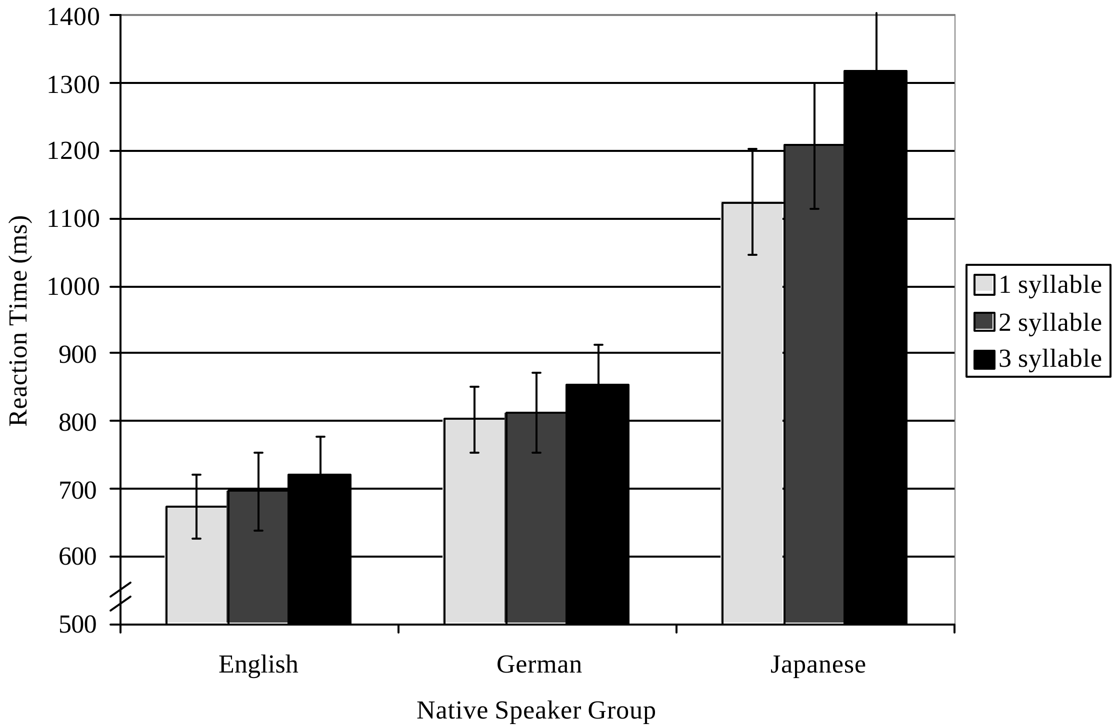


Figure 3.10. Mean RT (ms) by syllable length and language group for nonwords.

Post hoc analyses found that there was a main effect of syllable length because over all three groups, 1 and 2 syllable words were responded to significantly faster than 3 syllable words. However this main effect was qualified by a significant interaction with language group. As shown in the earlier analyses, this interaction arose because the syllable length effect for nonwords for each of these three groups was slightly different. For native speakers of German, there were no significant differences in response time for any of the three syllable conditions. For native speakers of English, there was a hint of a syllable length effect but for this group three syllable words were only marginally slower than 1 syllable words. However for the native Japanese group, there were significant differences across syllable conditions in that three syllable words were responded to significantly slower than 1 syllable words.

There was also a significant main effect for language group. This main effect was due to significant reaction time differences across all three groups. Not surprisingly, the native English speaker group was significantly faster than both the native German (t_{69} = 2.69, $p < .01$) and native Japanese speaker group (t_{49} = 4.67, $p < .001$). In addition, the native German group was faster than the native Japanese group (t_{46} = 3.12, $p < .004$).

Nonword Accuracy. The findings were somewhat different for accuracy. A 3 (syllable length) X 3 (language group) ANOVA was used to analyze the arcsin transformed accuracy data for nonwords. This analysis did not find overall difference for

Table 3.15

Mean accuracy (out of 100%) for syllable type by language group

Language group	1 syllable	2 syllable	3 syllable
English	97.4	98.1	93.6
	(.07)	(.06)	(.06)
German	95.4	97.2	96.3
	(.07)	(.06)	(.06)
Japanese	96.2	94.1	95.8
	(.18)	(.15)	(.15)

Standard errors in parentheses

syllable length by participants (*F1*(2,164) = 1.18, *MSE* = .08, $p < .4$), or items (*F2*(2,69) = .61, *MSE* = .24, $p < .6$), or an overall difference between language groups (*F1*(2,82) = .50, *MSE* = .11, $p < .7$; *F2*(2,138) = .22, *MSE* = .07, $p < .8$). There was however an interaction between syllable length and language group that was significant by participants (*F1*(4,164) = 2.65, *MSE* = .08, $p < .04$), but marginal by items (*F2*(4,138) = 2.35, *MSE* = .07, $p < .06$). As revealed in the earlier analyses, this interaction was due to significant higher accuracy for 1 syllable words than 3 syllable words for the native English speaker group, and the lack of an effect for syllable length for both the native German and native Japanese speaker groups.

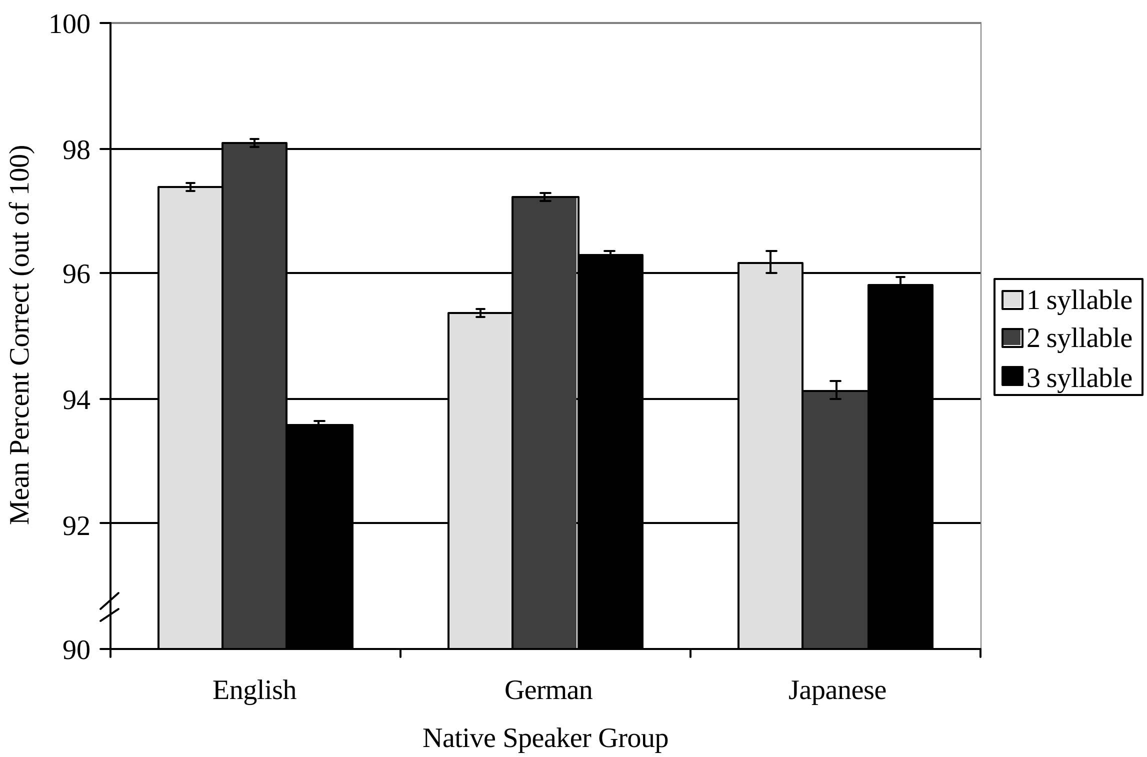


Figure 3.11. Mean accuracy by syllable length and language group for nonwords.

The results suggest that as processing is unfolding during the decision process for these nonwords, native Japanese speakers are affected by the overall length of the word more than native German speakers and slightly more than native speakers of English. Given the high consistency of the alphabetic form of Japanese, this result is perhaps not surprising and suggests that this language group seems to be relying on more serial first language processing strategies. Perhaps the more surprising result is that the native German speakers showed less of a syllable effect than native English speakers. Because empirical research has generally found that native German speakers demonstrate length effects in their first language, if these speakers were using the processing strategies from their L1, an expected length effect should emerge. However this native speaker group showed no differences across syllable length for RT or accuracy. Although this may suggest that this speaker group is using English processing strategies, differences between this groups and the native English group suggest that the strategies are not purely native-like. However overall the strategies used by the native German group are extremely similar.

Word Reaction Time. The word data for these groups however pattern in a slightly different way than the nonword data. For these stimuli, processing across the three language groups is more distinct. A 3 (syllable length) X 3 (language group) ANOVA was used to examine the processing differences between native English, German, and Japanese speakers for real English words. This analysis revealed a main effect for syllable length by participants ($F1(2,164) = 53.84$, $MSE = 7253.1$, $p < .001$), and items ($F2(2,69) = 12.55$, $MSE = 31872.1$, $p < .001$). This analysis also found a significant main effect for language group by participants ($F1(2,84) = 21.16$, $MSE =$

61836.0, $p < .05$), and items ($F2(2,138) = 10.95$, $MSE = 10936.3$, $p < .001$). The main effect for language group for the word data had the same overall pattern as nonword data.

Table 3.16

Mean RT (ms) for syllable type by language group for words

Language group	1 syllable	2 syllable	3 syllable
English	539	573	629
	(18.9)	(25.4)	(32.6)
German	644	663	708
	(19.7)	(26.5)	(34.0)
Japanese	708	909	1000
	(30.7)	(41.2)	(53.0)

Standard errors in parentheses

Native English speakers had faster response times than both the native German ($t_{69} = 3.66$, $p < .002$) and native Japanese ($t_{49} = 5.98$, $p < .001$) speakers. Once again, native German speakers had faster response times than native Japanese speakers ($t_{46} = 3.60$, $p < .003$).

The main effect for syllable length was due to significant differences between all three syllable groups. Post hoc analyses found that 1 syllable words had significantly faster response times than both 2 syllable words ($t_{84} = 4.39$, $p < .001$) and 3 syllable words ($t_{84} = 6.31$, $p < .001$). Also, 2 syllable words were recognized significantly faster then 3 syllable words ($t_{84} = 4.67$, $p < .001$). However this main effect was qualified by a significant interaction with language group. As found in earlier analyses, this interaction is due to three different response patterns across the three language groups. Differences across all three syllable lengths were significant for the native English speakers. For native German speakers, 1 and 2 syllable words had faster response times than 3 syllable words and for native Japanese speakers, 1 syllable words were faster than both 2 and 3 syllable words. These results go against the general prediction that if all three groups are relying on first language processing strategies then native speakers of English should show the least amount of a syllable effect, followed by the native German and then native Japanese speakers. Looking at the RT data suggests that the Japanese native speakers do have larger differences across number of syllables than the other two groups. However this effect may not have reached significance due to the small number of Japanese participants (n=14) and higher variability in proficiency than the other two groups.

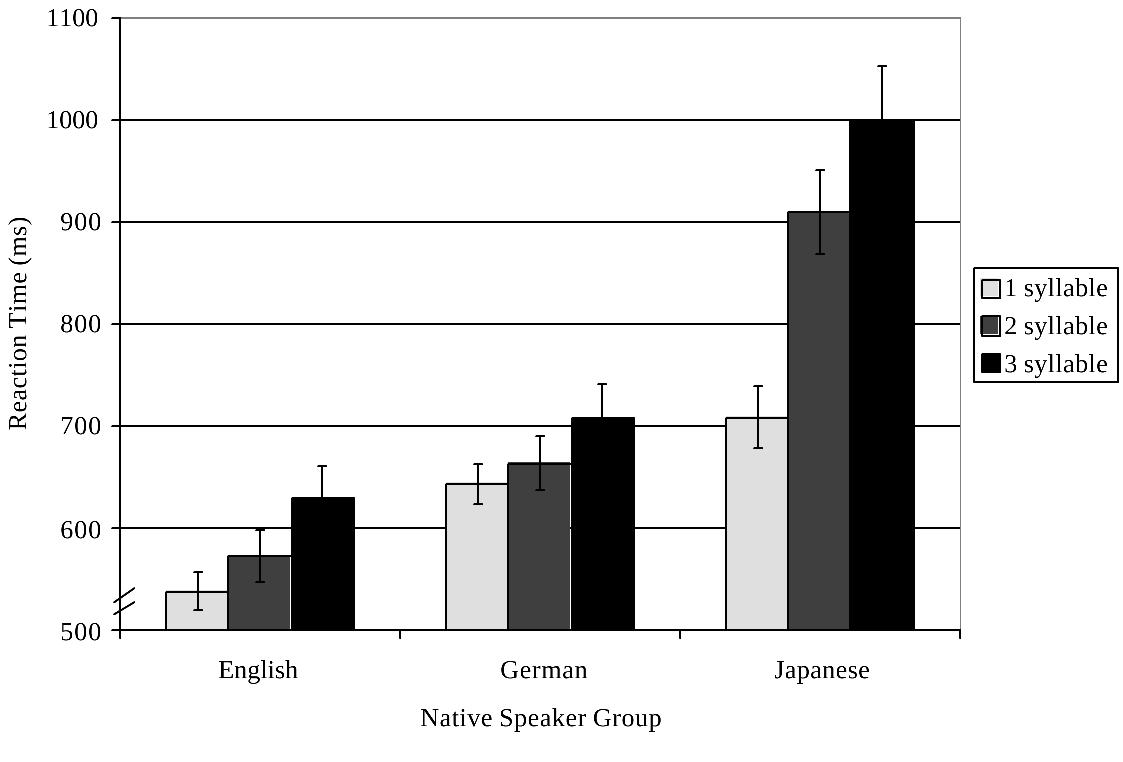


Figure 3.12. Mean RT (ms) by syllable length and language group for words.

Word Accuracy. Once again accuracy did not follow the same pattern as the RT data for the word stimuli. A 3 (syllable length) X 3 (language group) ANOVA was utilized to examine the relation between syllable length for the three language groups.

Table 3.17

Mean accuracy (out of 100%) for syllable type by language group for words

Language group	1 syllable	2 syllable	3 syllable
English	98.9 (.04)	98.8 (.04)	95.8 (.04)
German	93.6 (.04)	92.0 (.05)	98.3 (.05)
Japanese	99.4 (.10)	95.9 (.11)	96.3 (.12)

Standard errors in parentheses

This analysis revealed a main effect of syllable length by participants (*F1*(2,164) = 3.82, *MSE* = .05, $p < .03$), but not items (*F2*(2,69) = 1.38, *MSE* = .18, $p < .3$). And a main effect of language group by participants (*F1*(2,82) = 9.80, *MSE* = .08, $p < .001$), but not items (*F2*(2,138) = .43, *MSE* = .06, $p < .7$). Post hoc analyses revealed that the main effect of syllable length for the participant analysis was due to marginally more accurate responses to 1 syllable words compared to 2 syllable words ($t_{84} = 2.17$, $p < .1$). For language groups, native English speakers were more accurate than native German speakers ($t_{69} = 4.74$, $p < .001$), and unlike the nonword analysis, native Japanese participants were more accurate than native German speakers ($t_{46} = 2.52$, $p < .05$).

These main effects were qualified by an interaction of syllable length and nonword type that was significant by participants (*F1*(2,164) = 14.35, *MSE* = .05, *p* <

.001), but not items (*F2*(4,138) = 1.66, *MSE* = .06, *p* < .2). This interaction was driven by three different accuracy patterns for each language group. As mentioned earlier, there were no significant accuracy differences across syllable conditions for native English speakers. For native German speakers, 3 syllable words were more accurate than 1 and 2 syllable words, and for native Japanese speakers 1 syllable words were had higher accuracy rates than two syllable words.

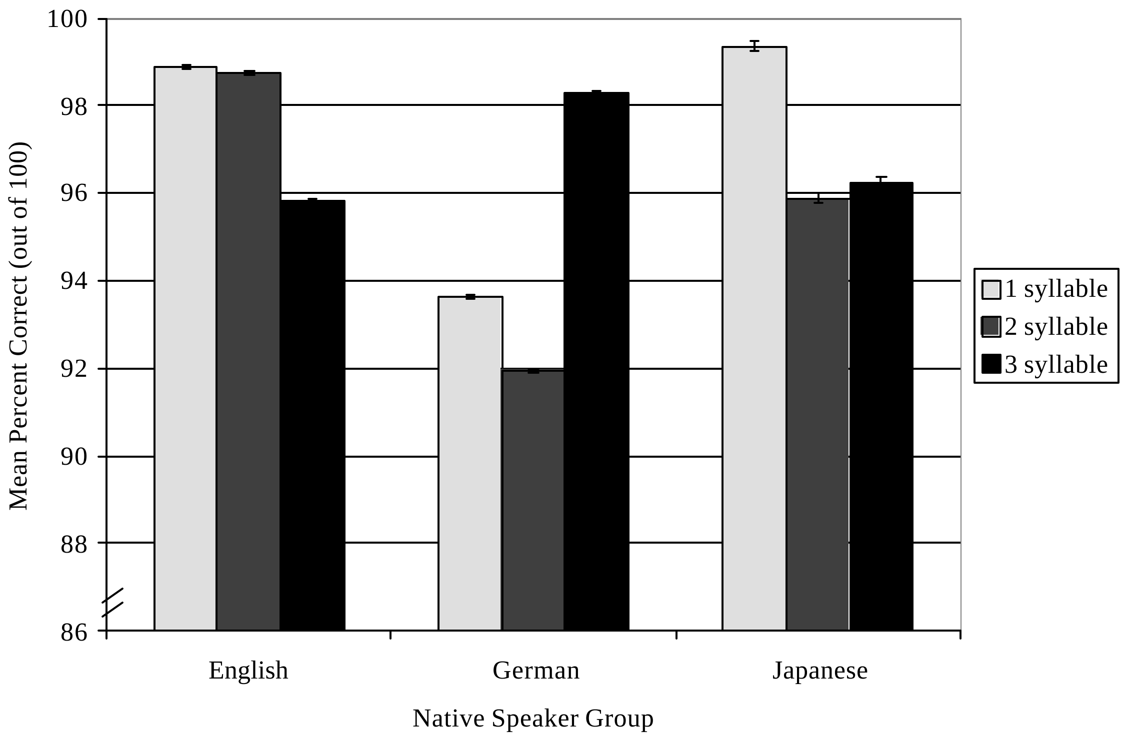


Figure 3.13. Mean accuracy by syllable length and language group for words.

For native English speakers, the word accuracy data closely mirrors the nonword data. For nonwords, three syllable words had lower accuracy because of the PSH nonword condition. It is possible that because the phonology was fully processed for both words and nonwords in the 3 syllable condition, and that word-like phonology was not a valid cue for whether or not a stimulus was a word, native English speaking subjects sometimes made no responses to word items. For the other two language groups, the word and nonword accuracy data are relatively dissimilar. Unlike the native English participants, native German participants actually show an accuracy advantage for 3 syllable words. It is possible that this advantage could arise from the overall familiarity of these words to nonnative speakers, however if this were true, the pattern would be similar for native Japanese speakers. However, this is not the case because the native Japanese group is most accurate for 1 syllable words. Therefore for these two groups, it is unclear how word accuracy informs us about L2 lexical parsing preferences in these two groups.

Analysis across language groups: Nonword type

Reaction Time. The second variable of interest across the three language groups was nonword type. To examine differences in nonword type across native speakers of English, German, and Japanese, a 3 (nonword type) X 3 (language group) ANOVA was

Table 3.18

Mean RT (ms) for nonword type by language group

Nonword type	O-P-	O+P+	PSH
English	675	723	692
	(53.6)	(54.0)	(49.2)
German	799	846	825
	(55.9)	(56.3)	(51.3)
Japanese	1186	1240	1225
	(87.1)	(87.8)	(79.9)

Standard errors in parentheses

conducted. This analysis found a main effect for language group by participants ($F1(2,82) = 13.98$, $MSE = 295822$, $p < .001$), and items ($F2(2,138) = 352.97$, $MSE = 14522.1$, $p < .001$). This main effect was due to significant differences in overall reaction time between the native English and Japanese groups, and also significant differences between the native German and Japanese groups (see syllable analysis for post hoc results). This analysis also revealed a main effect for nonword type by participants ($F1(2,164) = 11.24$, $MSE = 3834.4$, $p < .001$), but not by items ($F2(2,69) = 1.82$, $MSE = 28970.7$, $p < .2$). The main effect for nonword type by participants was due to significant differences between O-P- and O+P+ nonwords ($t_{84} = 6.52$, $p < .001$), and O+P+ nonwords and PSHs ($t_{84} = 2.62$, $p < .04$), and a marginal difference between O+P+ and PSHs ($t_{84} = 2.17$, $p < .1$). On average, response times to O+P+ nonwords were 48 ms slower than O-P- nonwords, and 24 ms slower than PSHs.

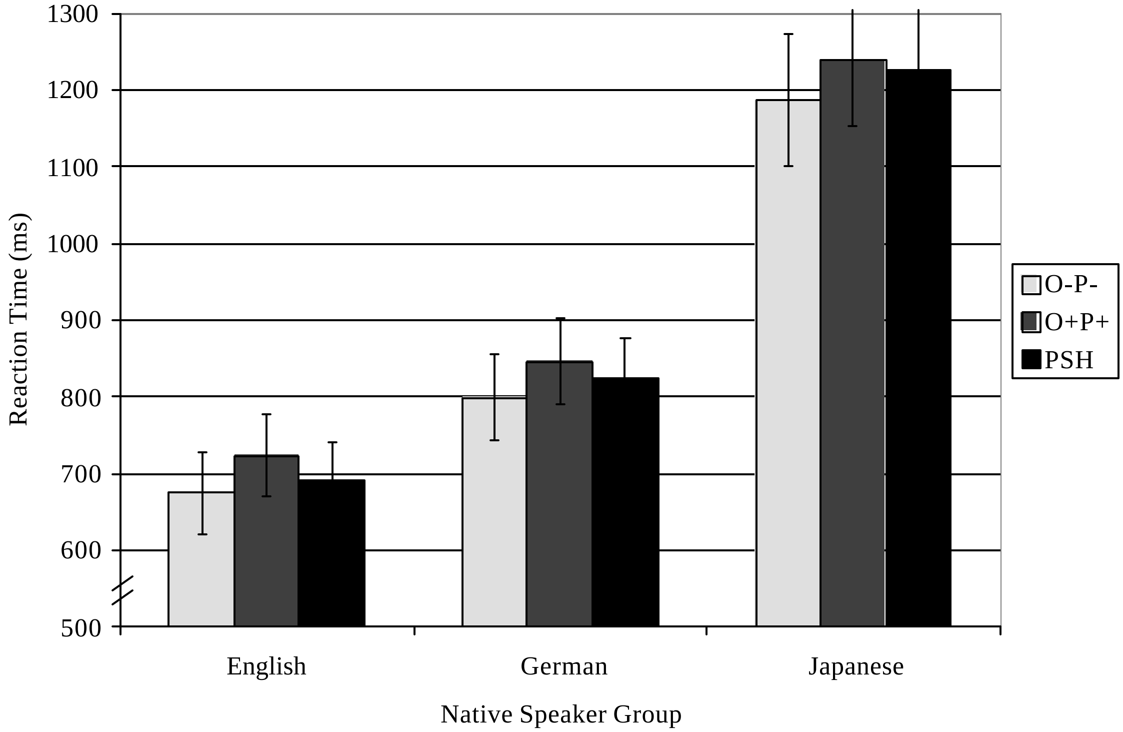


Figure 3.14. Mean RT by nonword type and language group.

Perhaps the most surprising result was that the interaction between language group and nonword type was not significant by participants ($F1(2,164) = .18$, $MSE = 3834.4$, $p < 1.0$), or items ($F2(2,69) = .19$, $MSE = 20184.7$, $p < .9$). The failure to find this interaction demonstrates that the patterns of response times over the three nonword types for the language groups were not significantly different from each other. The similarities between these three groups are shown in Figure 3.17. For all three language groups, O-P- nonwords had the fastest average response times followed by PSHs and then O+P+ nonwords.

Accuracy. The accuracy results over nonword type and language group closely match those results found for reaction time. To examine accuracy, a 3 (nonword type) X 3 (language group) ANOVA was conducted on arcsin transformed data. This analysis

Table 3.19

Mean accuracy (out of 100%) for nonword type by language group

Nonword type	O-P-	O+P+	PSH
English	98.2	95.7	95.5
	(.05)	(.07)	(.04)
German	98.4	95.7	94.2
	(.05)	(.07)	(.04)
Japanese	98.2	95.3	91.6
	(.12)	(.18)	(.10)

Standard errors in parentheses

failed to find a main effect for language group by participants (*F1*(2,82) = .50, *MSE* = .11, $p < .7$), or items (*F2*(2,138) = .22, *MSE* = .07, $p < .9$). There was a main effect for nonword type by participants (*F1*(2,164) = 17.74, *MSE* = .06, $p < .001$), and by items (*F2*(2,69) = 3.79, *MSE* = .22, $p < .03$). This main effect was due to higher accuracy for the O-P- condition over both the O+P+ condition ($t_{84} = 4.04$, $p < .001$) and the PSH condition ($t_{84} = 6.42$, $p < .001$).

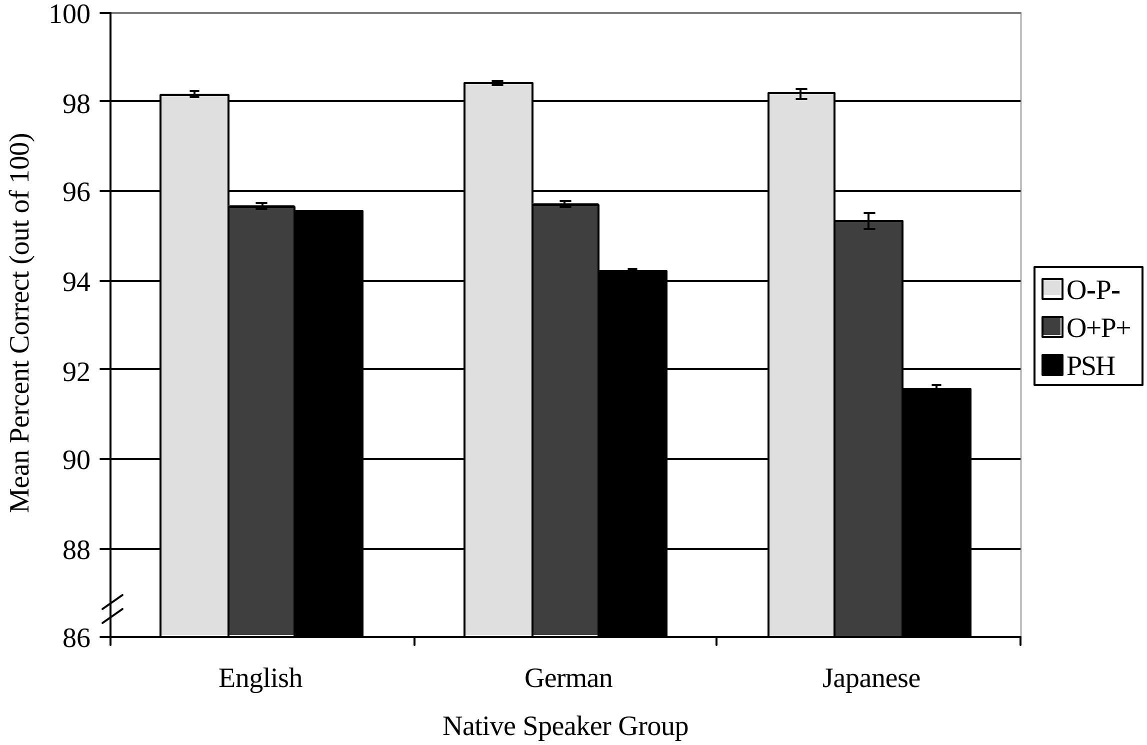


Figure 3.15. Mean RT by nonword type and language group.

Overall the lack of an interaction for language group and nonword type for both reaction time and accuracy was somewhat surprising. Given the greater consistency of German and Japanese, these two groups should have shown a larger PSH effect than the native English group if these participants were relying mainly on L1 processing strategies. The findings from these analyses suggest several things. First, slower reaction times for the O+P+ nonwords suggests that the nonnative speakers in this study know a number of the actual word neighbors for these nonwords that is similar to the native English speakers. These nonwords were responded to most slowly due to the high activation of word competitors. For error rates, PSH items were problematic for all three language groups suggesting that all groups were able to access the phonology of these nonwords and have it interfere to some degree.

Native German analysis by memory span

As mentioned at the beginning of this chapter, at the end of this experiment, the native German participants were presented with a German language version of the Waters and Caplan (1996) reading span task. Although there were no differences between high and low memory span participants for the native English speakers, the analysis by memory span for the native German participants did reveal some differences in performance between the two memory span groups. Memory span interacted with both syllable and length for reaction time to nonwords by participants ($F1(4,128)=3.20$, $MSE = 5117.0$, $p < .02$), and with syllable for reaction time to words ($F1(2,64)=4.66$, $MSE = 2522.8$, $p < .02$). For nonwords, after Bonferroni correction, there were no significant differences between any of the nonword conditions for the low memory span participants. However for the high memory span participants, one comparison did reach significance. For 2 syllable words, O-P- nonwords were significantly faster than O+P+ nonwords by an average of 76 ms ($t_{21} = 5.07$, $p < .001$).

Although the exact pattern of results for the high span German participants is not identical to the native English speakers, the general pattern is the same. There are consistent differences between O-P- and O+P+ nonwords, however the differences between these two types of nonwords and PSHs are not as consistent. Overall these findings suggest that the high span participants are better able to suppress L1 processing and use L2 processing strategies.

General Discussion

Overall the findings from this study provide some evidence that nonnative speakers are able to process L2 words and nonwords like native speakers, especially those participants with high working memory capacity. There were no differences between the three groups in the overall pattern for different nonword types, suggesting that across these three groups, similar orthographic and phonological information was used during processing. However these similarities across the nonword groups were marked by differences across numbers of syllables. For syllable length, there were greater similarities in processing between the relatively homogeneous native German speaker group and native English group. These groups seem to show greater similarity in processing 1 and 2 syllable words than 3 syllable words. In contrast, Japanese participants seem to incrementally increase response time across the three syllable conditions.

The differences in response times to syllable length across the three groups do seem to reflect the differences in consistency across the three native languages. The presence of both a lack of differences between groups for nonword type and differences that reflect L1 processing for syllable length support a bilingual version of the dual route model where L2 words are added within the lexical route however the GPC rules guiding serial processing only include L1 rules (see figure 2.2, Chapter 2). For the less homogenous native Japanese speaker group, the pattern of data from the syllable analyses suggests that either L1 strategies are influencing word recognition in English or the overall longer response latencies are the source of the syllable effect. However for the proficient and relatively homogeneous native German speakers it is more difficult to determine whether the minor differences between this group and the native English speaker group are due to a systematic influence of L1 processing or whether these differences are due to the slightly slower response times of these participants. In the following chapter, a cluster lexical decision task will be used to address this question.

Chapter 4: Examining Parsing Through Clustered Lexical Decision

There have been a number of findings demonstrating that the orthographic rime may be an important unit of processing in English. As mentioned in Chapter 1, a recent study by Ziegler and Perry (1998) suggests that the effect of neighborhood size in word recognition in English may be a consequence of reliance on the orthographic rime. In languages in which there is a more consistent mapping from spelling to sound and a corresponding reliance on smaller sublexical units during word recognition, the same neighborhood density effects are not observed.

The logic of examining the orthographic rime has its foundation in a number of experiments and statistical analyses of the English language conducted by Treiman and her colleagues (Kessler & Treiman, 1997; Treiman & Chafetz, 1987; Treiman et al., 1995). These studies have shown that the orthographic rime is important in English because it has a fairly consistent spelling-to-sound mapping compared to smaller sublexical units. As noted earlier, similar studies with orthographically shallow languages, such as Dutch, have produced different results (e.g., Martensen et al., 2000), showing greater reliance on smaller sublexical units.

The present experiment is an adaptation of Treiman and Chafetz (1987). The logic of that study was to investigate the importance of the orthographic rime in English by asking how word recognition performance is affected when the form of the presentation makes the hypothesized parsing strategy difficult. Treiman and Chafetz used a technique in which they visually divided up words either by onset and rime, consistent with the preferred parsing, or after the first vowel, a division that was hypothesized to violate the preferred parsing. In their experiment, words were divided by placing two slashes either between the onset and the rime (e.g., PR//INT) or after the first vowel (e.g., PRI//NT). The basic result was that participants were faster to recognize words when the slashes fell between the onset and rime than when the slashes fell after the first vowel. As mentioned earlier, a criticism of the Treiman and Chafetz study is that for some of the stimuli that had the onset-rime division, the rime formed a real word in English (e.g., BL//IMP). It is therefore possible that the benefit in this condition was due to the recognition of the rime word and not to the coincidence of the break with the preferred parse.

The goal of the experiment reported here was to create a new method for examining lexical parsing by adapting the Treiman and Chafetz (1987) task so that the parsing preferences of native and non-native speakers of English could be compared.

Description of the Task

In creating the new method the first aim was to ensure that participants were responding to a complete and unbroken string of letters and minimize the possibility that participants were making judgments on the rime and not the entire string. To meet this goal, the decision was made to first present participants with either the onset or the onset and first vowel, and then to present the string in its entirety. Using this presentation sequence, two different versions of the task were created. In one version, participants were presented with either an onset-rime parsed stimulus, the onset followed by a blank space, or a nononset parsed stimulus, the onset plus the first vowel followed by a blank space. After a set period of time, the remainder of the string "filled in" and participants made a response as to whether or not the presented string was a word (see Figure 4.1).

In a second version of the task, instead of blank space following the first word segment, the space was filled with three asterisks if the first segment was an onset and two asterisks if the first segment was the onset plus the first vowel. After a specified amount of time, the remainder of the word was revealed in the space once held by the asterisks (see Figure 2.4).

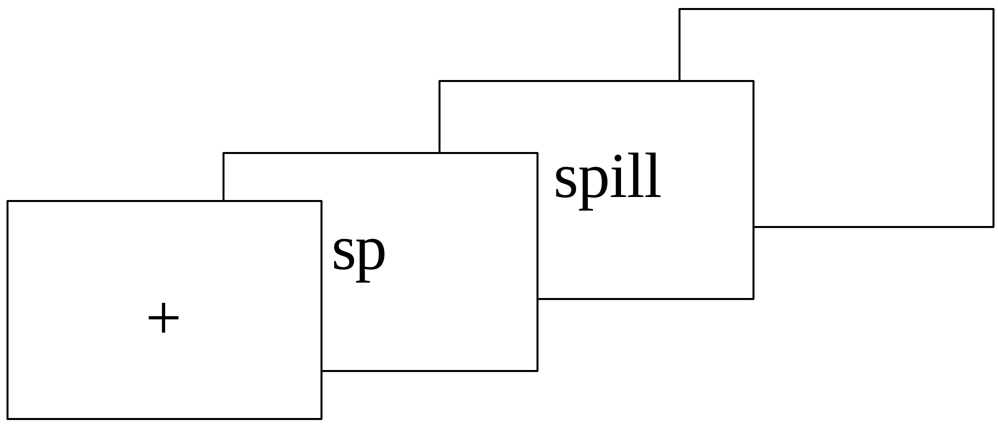


Figure 4.1. Events within the cluster lexical decision task in Experiment 2.

In each method of presentation, there were three delay conditions (SOAs of 50, 250, and 500ms) to determine whether violations of parsing preferences are more disruptive early or late in processing.

The lexical decision performance of two language groups was then compared: native English speakers and native German speakers who spoke English as their L2. A first goal was to determine whether native English speakers produce the same results on this variant of the task as those reported by Treiman and Chafetz (1987). The pattern of results for the native English speakers was then compared to those for the native German speakers who completed the task in English. If native German speakers transfer the preferred strategy associated with their L1 to parsing in the L2, the reliance on smaller unit processing strategies from German should make processing difficulty equivalent for both onset-rime parsed stimuli (e.g., BL BLANK) and nononset parsed stimuli (e.g., BLA BLANK).

Method

Participants

Two groups of participants were recruited for this study: native English speakers (n=140) and German-English bilinguals (n=54). The native English speakers were recruited from the psychology department subject pool and were given course credit for their participation. The German-English bilinguals were recruited from the University of Muenster and the Max Planck Institute of Cognitive Neuroscience in Leipzig and were paid for their participation.

Materials

To determine whether the method used in this experiment would produce findings analogous to those reported by Treiman and Chafetz (1987), the same materials were used as in the previous study. All words and nonwords were five letters long and the pronunciation of these letter strings contained either three phonemes (e.g., THING) or five phonemes (e.g., CREST). There were a total of 24 words in the 3 phoneme condition and 36 in the 5 phoneme condition. Words were blocked by the number of phonemes and all participants were presented with the block of 5 phonemes first, the longer of the two blocks.

Two versions of the materials were constructed. In one version, half of the words and nonwords were randomly assigned to be split after the onset and the remaining half of the words and nonwords were split after the first vowel. In the other version, the assignment was counterbalanced to create a mirror-image condition.

Procedure

Participants completed the task one at a time while seated at the computer. Native English speakers were assigned to one of 12 different scripts (2 method of presentation X 3 SOA X 2 word division). On each trial, a fixation appeared on the screen for 500 ms followed by the first letter cluster. For half the participants, this first letter cluster was followed by asterisks and for the other half, only blank space. The first letter cluster was then followed by the presentation of the entire letter string either at 50, 250, or 500 ms. The SOA was a between-participant factor so that SOA was constant for any given participant. Individuals were instructed to press a key marked “1” if the presented string was a nonword, and a key marked “2” if the string was a real English word. Each participant was assigned to one of the two versions of the materials. For the native German participants, the procedure was identical with a single exception. Because the pool of German participants was smaller than that of native English speaking participants, only the short (50 ms) and long (500 ms) SOAs were used.

Results and Discussion

As in Experiment 1, each language group was first analyzed separately to examine how word division affected word recognition. First, the results were collapsed across the three SOA conditions and examined by method of presentation to determine whether the findings from this experiment differ from those reported by Treiman and Chafetz (1987), and whether method of presentation (space vs. asterisks) affected how the words and nonwords were processed. The results will then be analyzed to determine whether SOA affected the parsing strategies of the participants. In the final analyses, both language groups will be compared to determine whether they differ. Because only two SOA conditions were used with the native German participants, only native English data from the 50 and 500 ms SOA conditions will be included.

Exclusion of Participants

At the end of the experiment, participants completed a language history questionnaire. Participants were excluded if they were not native English or German speakers or if another language was spoken in the home. For the native English speaker group, 13 participants were removed from the analysis because they did not meet the language criteria, and for the native German speaker group, three participants were removed because they did not meet the language criteria.

In addition to the language criteria, those participants who did not meet a set of accuracy criteria were also excluded. Because participants overall had lower accuracy on this task than in Experiment 1 they were excluded only if their accuracy for words and/or nonwords was below 70%. (Recall that the materials in Experiment 1 were taken from a study that had been designed for use with children.) For the native English speaker group, 11 participants were excluded due to low accuracy, and for the native German group, 14 participants failed to meet these accuracy criteria. After these exclusions, 116 participants remained in the native English group, and 37 in the native German group.

Native English speaker group

Reaction time for words by phoneme and parse

In this first analysis, reaction time to words was examined as a function of type of parse (onset-rime vs. nononset) and number of phonemes (3 vs. 5). The data were also grouped by method of presentation (space vs. asterisks).

Table 4.1

Reaction time (ms) as a function of type of parse and number of phonemes for words in Experiment 2 and Treiman and Chafetz (1987)

Method	Number of phonemes	Nononset Parse	Onset-rime parse	Parse difference
Space	3	718 (25.6)	749 (29.0)	-31
	5	765 (28.4)	780 (29.6)	-15
Asterisks	3	770 (23.2)	793 (26.3)	-23
	5	797 (25.8)	827 (26.9)	-30
Slashes (Treiman and Chafetz, 1987)	3	802	745	+57
	5	892	849	+43

Standard errors in parentheses

To examine the differences between onset-rime and nononset parsed words, an ANOVA was conducted with participants as a random factor with number of phonemes (3 or 5) X type of parse (onset-rime or nononset). This analysis was conducted separately for the blank space method of presentation and the asterisks method of presentation. In this analysis number of phonemes and type or parse were within participant factors. For the item analysis, type of parse was treated as a within-item factor and number of phonemes became a between-item factor.

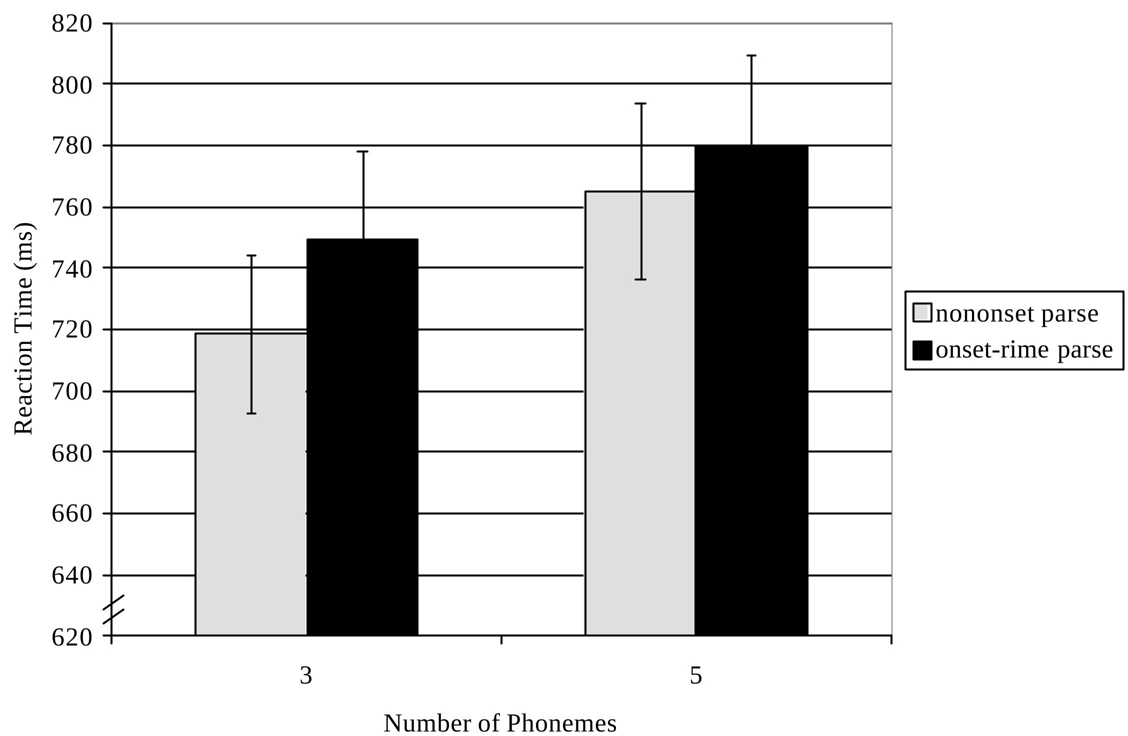


Figure 4.2. Reaction time (ms) for type of parse by number of phonemes for the space method of presentation.

For the space method of presentation, the analysis revealed a main effect of parse by participants (*F1*(1,52)=6.55, *MSE* = 5414.5, $p < .02$) but not items (*F2*(1,34)=0.83, *MSE* = 3119.2, $p < .4$). Surprisingly, this main effect in the participant analysis was due to significantly faster response times for nononset parsed stimuli by an average of 23 ms, a difference in the direction opposite to what was predicted. In the participant analysis there was also a main effect of number of phonemes ($F1(1,52)=6.46$, $MSE = 6417.5$, $p < .02$), however this effect was also not significant by items ($F2(1,34)= 0.04$, $MSE = 11774.2$, $p < .9$). The main effect of phonemes for participants was due to an average faster response time of 39 ms for 3 phoneme words over 5 phoneme words. Number of phonemes and type of parse failed to interact for both the participant analysis ($F1(1,52)=1.55$, $MSE = 2237.3$, $p < .3$) and the item analysis ($F2(1,34)= 0.04$, $MSE = 3119.2$, $p < .09$).

The results for the asterisks method of presentation were almost identical. This analysis revealed a main effect of type of parse in both the participant analysis ($F1(1,63)=16.66$, $MSE = 5094.3$, $p < .001$) and the item analysis ($F2(1,34)= 5.59$, $MSE = 2668.1$, $p < .03$). As in the space method analysis, this main effect of parse was due to faster response times for nononset parsed words by an average of 26 ms.

This analysis also revealed a main effect of number of phonemes by participants ($F1(1,63)=9.47$, $MSE = 8175.3$, $p < .004$) but not by items ($F2(1,34)= 1.08$, $MSE = 10036.6$, $p < .4$). Once again, the main effect by participants was due to faster response times for 3 phoneme words over 5 phoneme words by an average response time of 36 ms. As in the space method analysis, number of phonemes and type of parse did not interact by participants ($F1(1,63)=0.13$, $MSE = 4802.7$, $p < .8$) or by items ($F2(1,34)= 0.01$, $MSE = 2668.1$, $p < 1.0$).

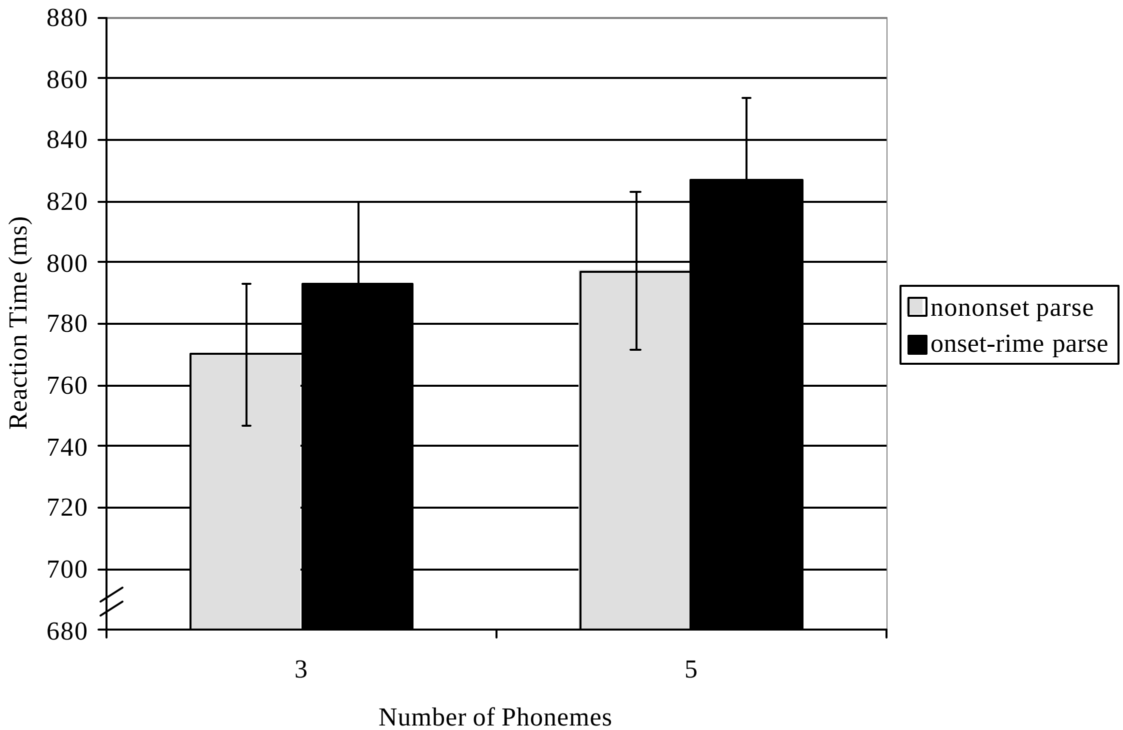


Figure 4.3. Reaction time (ms) for type of parse by number of phonemes for the asterisks method of presentation.

The results from these two sets of analyses are in the opposite direction of those reported by Treiman and Chafetz (1987). In their experiment they found an average processing advantage of 50 ms for onset-rime parsed words over nononset parsed words.

There are a number of reasons why the results from Experiment 2 differ from those of Treiman and Chafetz (1987). First, for the onset-rime parse condition in the present experiment, participants were making decisions to the entire string of letters after seeing the onset. It is possible that with this method of presentation, the rime unit became less salient than in Treiman and Chafetz's slash presentation. A second reason could lie

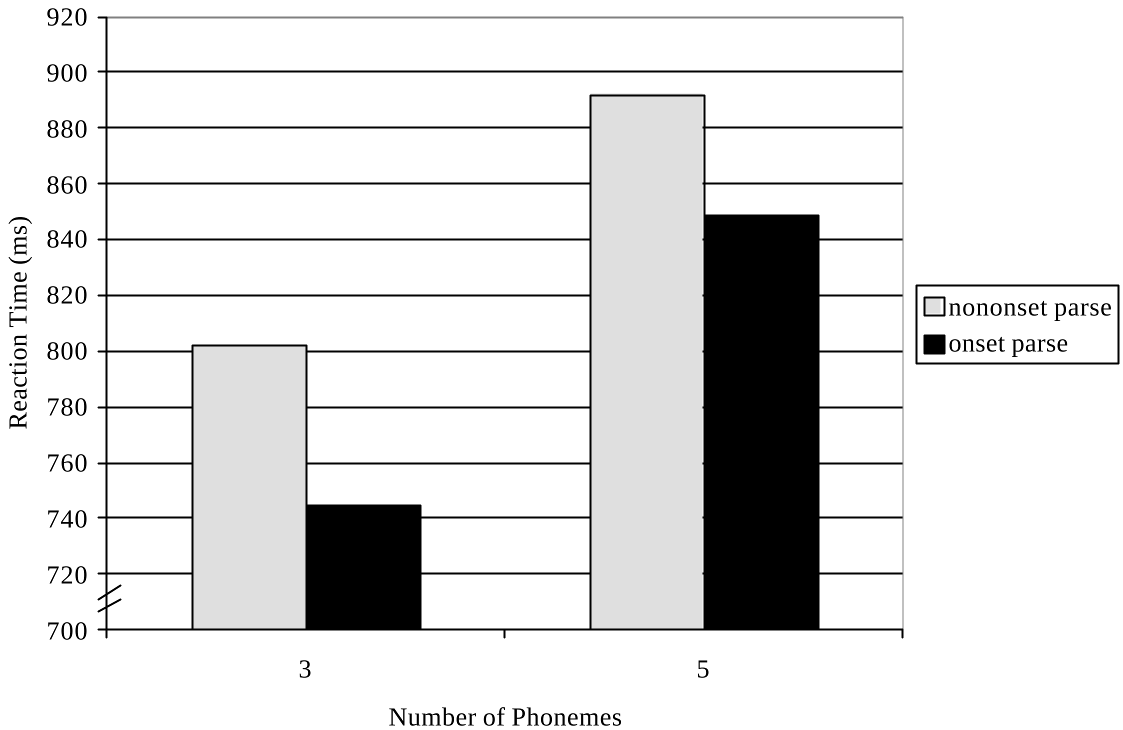


Figure 4.4. Reaction time (ms) for words by type of parse and number of phonemes for Treiman and Chafetz (1987).

within the decision speed for both groups. From Table 4.1, it is clear that on average the participants in Treiman and Chafetz's experiment were slower than the participants in Experiment 2. Because these participants were making decisions to words and nonwords divided by slashes, it is possible that the task induced a checking strategy that slowed lexical decisions. To determine whether there were greater similarities between the slower decision participants in Experiment 2 and Treiman and Chafetz's participants, the participants in the current experiment were divided into a fast and slow decision group. Participants whose mean response time was above 700 ms were placed into the slower decision group and those whose average response time was below 700 ms were placed into the faster decision group. This analysis by participants found a main effect of type of parse for both the faster participants (*F1*(1,73)=9.12, *MSE* = 1661.6, $p < .004$), and the slower participants (*F1*(1,42)=16.09, *MSE* = 10094.4, $p < .001$). The direction of the effect for these analyses was also the same such that participants responded faster to nononset parsed words than onset-rime parsed words. The finding that slower participants did not show the same effect as the participants in the Treiman and Chafetz study rules out the possibility that the differences are due to an extended processing period per se. This analysis does not, however, rule out the explanation that the rime may have been more salient in the Treiman and Chafetz experiment since for half of the trials, the rime was seen as a distinct unit. Alternative interpretations for the different findings between these two methods will be discussed at the end of this chapter.

Reaction time for nonwords by phoneme and parse

In this analysis, performance on nonwords was examined to determine whether the type of parse and number of phonemes affected how quickly participants were able to reject these items. Separate analyses again were conducted for the space and asterisks methods. Nonword data were analyzed using a 2 (type of parse) X 2 (number of phonemes) ANOVA. As in the word analyses, the findings for the nonwords diverged from those reported by Treiman and Chafetz (1987).

Table 4.2

Reaction time (ms) as a function of type of parse and number of phonemes for nonwords in Experiment 2 and Treiman and Chafetz (1987)

Method	Number of phonemes	Nononset parse	Onset-rime parse	Parse difference
Space	3	781	817	-36
		(32.9)	(30.8)	
	5	844	856	-12
		(37.8)	(37.0)	
Asterisks	3	873	892	-19
		(35.7)	(40.1)	
	5	898	914	-16
		(37.1)	(37.7)	
Slashes (Treiman and Chafetz)	3	976	951	+25
	5	1132	1126	+6

Standard errors in parentheses

Unlike Treiman and Chafetz (1987), for nonwords there were no effects of the parse condition for the space version of the task by participants ($F1(1,52)=2.62$, MSE = 12055.6, $p < .2$), or by items ($F2(1,34)=0.97$, MSE = 4263.4, $p < .4$). For the asterisks version of the task, there was no main effect of the parse condition by participants ($F1(1,63)=2.38$, MSE = 8157.9, $p < .2$), however there was a main effect by items ($F2(1,34)=6.95$, MSE = 4184.8, $p < .02$). This main effect by items was opposite to the predicted direction, with participants on average 44 ms faster to reject nononset parsed nonwords than onset-rime parsed nonwords.

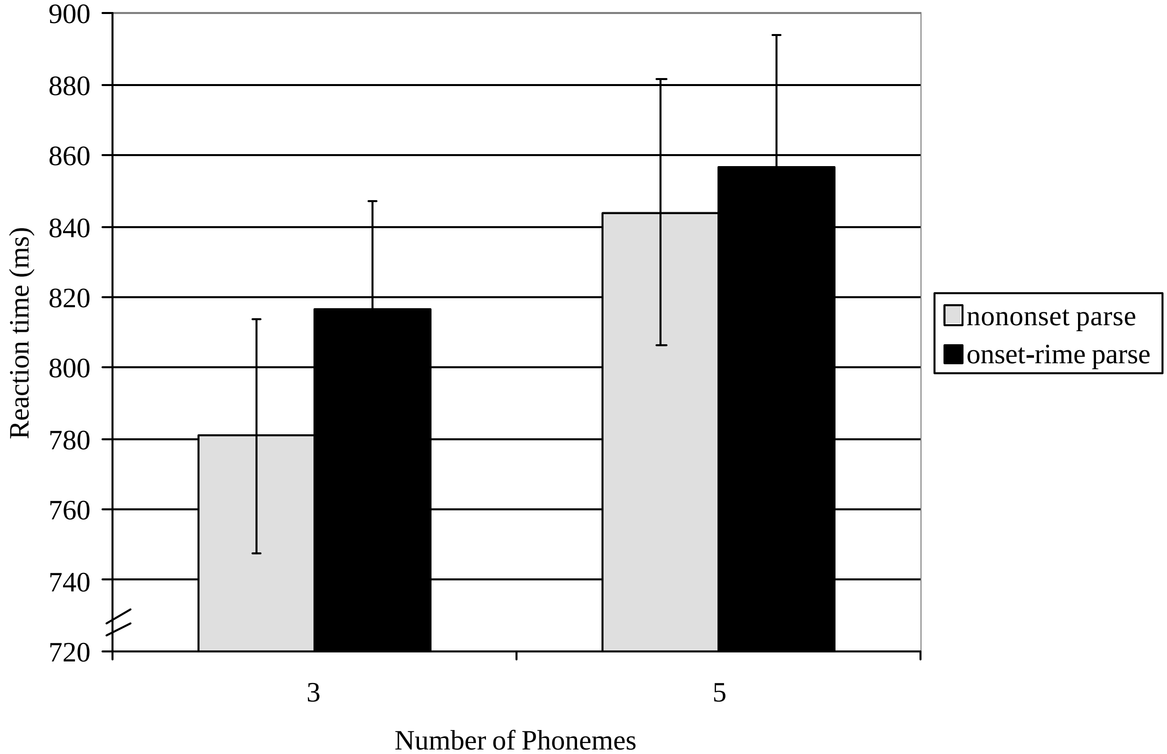


Figure 4.5. Reaction time (ms) for type of parse and number of phonemes for nonwords for the space method of presentation.

This analysis also revealed significant differences between the two phoneme conditions. For the space method of presentation, 3 phoneme nonwords were rejected on average 49 ms faster than 5 phoneme nonwords. This difference was significant by participants ($F1(1,52)=8.43$, $MSE = 16832.5$, $p < .006$) but not items ($F2(1,34)=1.47$, $MSE = 25162.7$, $p < .3$). For the asterisks method, the effect of number of phonemes was marginal for participants ($F1(1,63)=3.54$, $MSE = 9812.3$, $p < .07$) and not significant by items ($F2(1,34)=0.68$, $MSE = 15980.5$, $p < .5$).

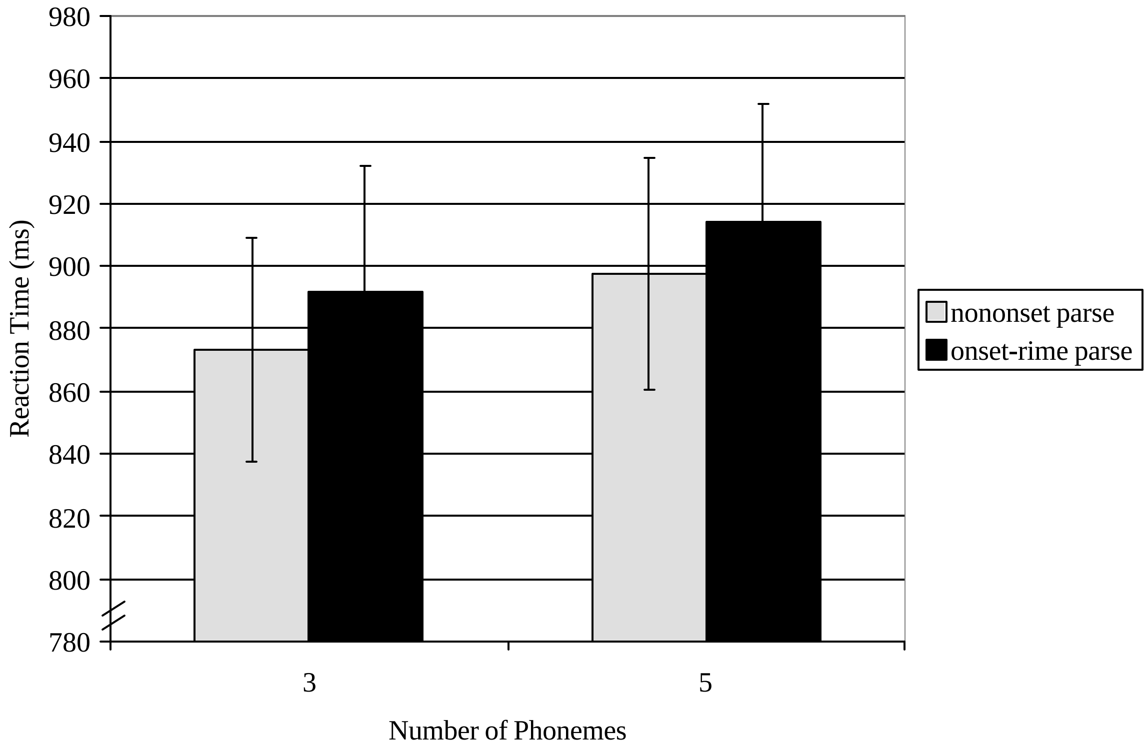


Figure 4.6. Reaction time (ms) for type of parse and number of phonemes for nonwords for the asterisks method of presentation.

Although these findings for phonemes seem somewhat divergent with Treiman and Chafetz's (1987) findings, their data do show some evidence that 3 phoneme nonwords were rejected faster than 5 phoneme nonwords although these differences failed to reach significance.

Overall, the findings of this experiment diverge from Treiman and Chafetz (1987). When participants were presented with the onset first followed by the rime, response times to both words and nonwords were actually slower than when participants were given the onset plus the vowel followed by the remainder of the word. One possible explanation for the observed differences is that the new methods used in this experiment had the effect of disrupting processing more for the onset-rime parsed stimuli than for the nononset rime parsed stimuli. If during processing the rime unit has more of an influence on processing than the onset, receiving the entirety of the rime unit later in the stream may be disruptive to processing.

Alternatively, when participants were presented with the nononset parse, they received three letters first instead of two. Having the extra information provided by that letter 50, 250, or 500 ms prior to the presentation of the remainder of the string may have provided a larger processing benefit than having the preferred parse. This point will be revisited at the end of this chapter.

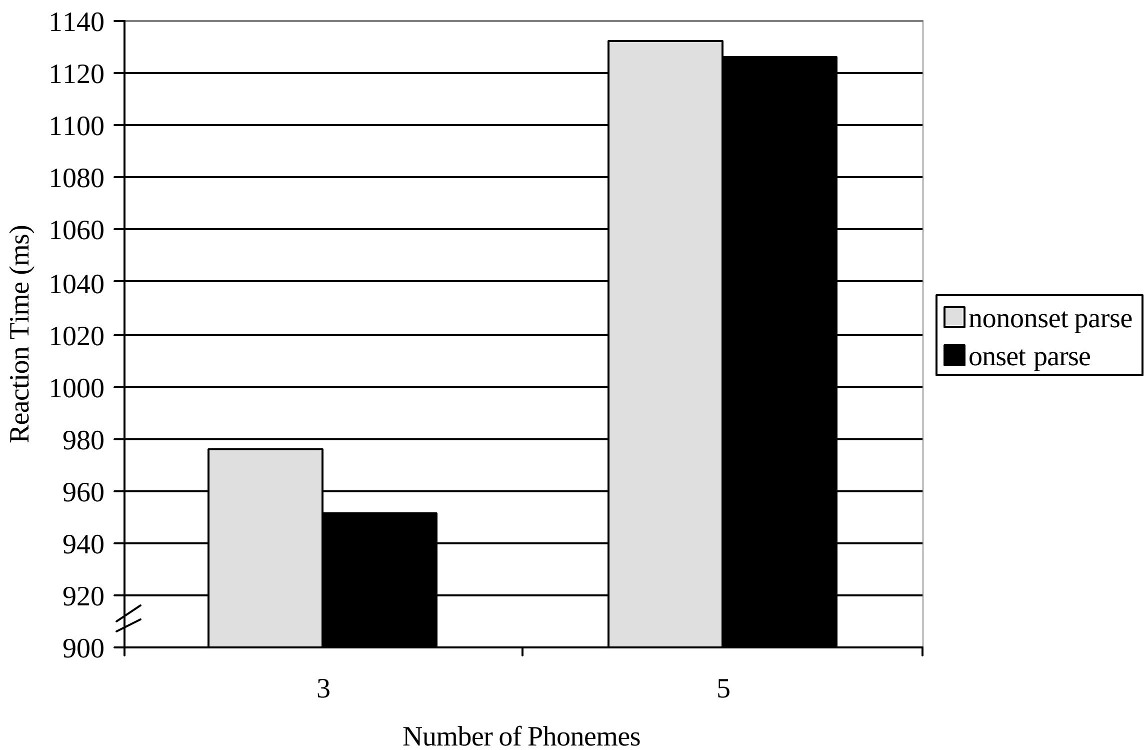


Figure 4.7. Reaction time (ms) for nonwords by type of parse and number of phonemes for Treiman and Chafetz (1987).

Although the findings from this experiment are at odds with those of Treiman and Chafetz (1987), the results across the two methods of presentation were fairly consistent. For both the space and asterisks manipulations, nononset parsed words were responded to faster than ones parsed by onset-rime, and 3 phoneme words were responded to faster than 5 phoneme words. In addition, average response times for these two methods did not differ from each other statistically for words ($t_{115} = 1.03$, $p < .4$) or nonwords ($t_{115} = 1.41$, $p < .2$).

Reaction Time comparison by SOA for words

The second manipulation in this study was the use of three different SOAs: 50, 250, and 500 ms. The SOA manipulation was included to determine whether processing was disrupted more at a short, medium, or long lag between the presentation of the first letter cluster and the remainder of the word. A differential effect of SOA on performance would have important implications for identifying the locus of parsing. For the participant analysis, number of phonemes and parse were within-participant variables and SOA was a between-participant variable. For the item analysis parse and SOA were within-item variables and number of phonemes was a between-item variable. For both analyses a 2 (number of phonemes) X 2 parse X 3 (SOA) ANOVA was conducted to examine the effects for these three variables.

Table 4.3

Reaction time (ms) as a function of type of parse and number of phonemes for words by SOA in Experiment 2

SOA (ms)	Number of phonemes	Nononset parse	Onset-rime parse	Parse difference
50	3	640	671	-31
		(24.0)	(28.0)	
	5	691	709	-18
		(24.1)	(27.2)	
250	3	635	672	-37
		(20.7)	(24.1)	
	5	677	707	-30
		(20.8)	(23.4)	
500	3	690	722	-32
		(21.0)	(24.4)	
	5	702	740	-38
		(21.1)	(23.7)	

Standard errors in parentheses

This analysis revealed a main effect of parse by participants, ($F1(1,114)=16.76$, $MSE = 5294.0$, $p < .001$) and by items ($F2(1,34)=4.91$, $MSE = 4571.5$, $p < .04$). Once again, this main effect by participants was due to faster response times to nononset parsed words than onset-rime parsed words by an average of 32 ms. The analysis also revealed a main effect of number of phonemes by participants ($F1(1,114)=16.76$, $MSE = 7295.7$, $p < .001$) but not items ($F2(1,34)=.39$, $MSE = 29577.0$, $p < .6$). As in the previous analyses, the main effect by participants was a result of faster response times for 3 phoneme words compared to 5 phoneme words by an average of 33 ms. There was no effect of SOA in the participant analysis ($F1(2,144)=1.18$, $MSE = 69555.2$, $p < .4$) however this effect was significant by items ($F2(2,68)=8.68$, $MSE = 2612.3$, $p < .001$). The effect by items was due significant response time differences between the 250 ms and 500 ms SOA conditions ($t_{35} = 4.22$, $p < .001$). On average, response times in the 250 ms condition were 35 ms faster than the 500 ms condition. No other paired comparisons reached significance.

Overall reaction time performance for native speakers of English was affected by how words were parsed, the number of phonemes, and also the time delay between the presentation of the first letter cluster and the remainder of the word. Once again, the results from this analysis diverge from those results reported by Treiman and Chafetz. Nononset rime parsed stimuli were responded to faster than onset-rime parsed stimuli, and this difference in parse was similar across the two phoneme conditions and three SOA conditions. Another interesting finding from this analysis was that response time for the 250 ms condition was significantly faster than the 500 ms condition and also slightly faster than the 50 ms condition (however this difference was not significant). This pattern of reaction time across these three SOAs suggests that processing was disrupted the least amount in the 250 ms condition. Although the difference was not significant, participants were somewhat slower in the 50 ms condition perhaps suggesting that they had not yet finished processing the first letter cluster. If this is the case, participants in the 500 ms condition would have finished processing the first letter cluster around 250 ms. By 500 ms it is possible that the target word and other words sharing that letter cluster had been activated therefore creating more competition and slower response times at this SOA.

Reaction Time comparison by SOA for nonwords

For nonwords the results were also similar across the three SOAs. There was a main effect for parse by participants (*F1*(1,114)=4.55, *MSE* = 9901.4, p < .04) and by items (*F2*(1,34)=7.72, *MSE* = 5357.0, p < .009). As in the word analysis, the main effect of parse was due to faster response times for the nononset parse condition by an average of 20 ms.

Table 4.4

Reaction time (ms) as a function of type of parse and number of phonemes for nonwords by SOA in Experiment 2

SOA (ms)	Number of phonemes	Nononset parse	Onset-rime parse	Parse difference
50	3	857	855	+2
		(47.7)	(50.3)	
	5	882	905	-13
		(51.2)	(51.1)	
250	3	806	839	-33
		(41.2)	(43.4)	
	5	863	859	+4
		(44.2)	(44.1)	
500	3	838	880	-42
		(41.6)	(43.9)	
	5	877	905	-28
		(44.7)	(44.6)	

Standard errors in parentheses

This analysis also revealed a main effect of number of phonemes by participants (*F1*(1,114)=11.37, *MSE* = 13296.7, p < .002) but not by items (*F2*(1,34)=1.17, *MSE* = 56613.2, p < .3). Similar to the word analysis, the main effect of number of phonemes by participants was due to a significant difference of 36 ms between the 3 phoneme condition and the 5 phoneme condition.

Finally there was an effect of SOA that was not significant by participants (*F1*(2,114)=0.21, *MSE* = 288921.9, p < .9) but was significant by items (*F2*(2,68)=3.57, *MSE* = 9480.2, p < .04). After Bonferroni correction, post hoc comparisons revealed that this main effect was due to marginal differences between the 50 ms and 250 ms conditions (t_{35} = 2.35, $p < .08$), and also between the 250 and 500 ms conditions (t_{35} = 2.23, $p < .1$). Once again, of these three conditions, the 250 ms SOA had the fastest average response time. This condition was 42 ms faster than the 50 ms SOA condition, and 25 ms faster than the 500 ms SOA condition.

In general, the reaction time data for type of parse and number of phonemes for the words and nonwords for this experiment were consistent across both methods of presentation and across SOAs. Unlike the results of Treiman and Chafetz (1987), words presented with an onset-rime parse had significantly slower response times. For phonemes, the results were more similar across words and nonwords. In most cases, 3 phoneme stimuli were responded to faster than 5 phoneme stimuli. Although part of this effect for words may have been due to the overall greater frequency of the 3 phoneme words compared to the 5 phoneme words, the presence of a somewhat smaller effect for the nonwords in this experiment suggests that number of phonemes did play a role in how quickly the stimuli were processed.

The analyses reported were performed separately to examine method of presentation and SOA in order to avoid spurious interactions that arose from having a large number of variables within the analysis. However when the omnibus analysis was conducted there was one variable that seemed to consistently interact with other experiment variables. In the participant analysis, although words were randomly assigned to the two different scripts (which determined how they were parsed), script interacted with a number of important variables. First, script interacted with type of parse (*F1*(1,105)=8.65, *MSE* = 5171.3, p < .005). This interaction occurred because the difference in response time for onset and nononset parsed words was significant for script A (t_{60} = 4.80, *p* < .001) but not for script B (t_{55} = 1.52, *p* < .2). On average, nononset parsed words were responded to 51 ms faster for script A but only 10 ms faster for script B. However the predicted pattern of faster response times for onset-rime parsed items was not obtained in either script.

Script also interacted with type of parse and number of phonemes (*F1*(1,105)=10.85, *MSE* = 3232.4, p < .002). This three-way interaction occurred because for script A, the faster responses for nononset parsed words were significant for both 5 phoneme words (t_{60} = 2.86, *p* < .006), and three phoneme words (t_{60} = 4.60, *p* < .001). For script B, the difference was only significant for the 5 phoneme words (t_{60} = 3.39, *p* < .002). For script A, the nononset parsed items were 72 ms faster for 3 phoneme words and 30 ms faster for 5 phoneme words. For script B, in the 5 phoneme condition, nononset parsed words has an average faster response time of 29 ms.

In general, these interactions seemed to arise because the effect of parse was more consistent for the division for words in script A. Because words were randomly assigned a location of parse for script A and then mirrored for script B, it is unclear why this interaction occurred.

Accuracy for words by phoneme and parse

As in the word analysis, accuracy was examined as a function of type of parse and number of phonemes. For this first analysis, the two methods of presentation, space and asterisks, were examined separately and compared to the results from Treiman and Chafetz (1987). Because the overall accuracy was lower and more variable than Experiment 1, an arcsin transform was not performed and percent correct was used as the dependent measure.

Table 4.5

Accuracy (out of 100%) as a function of type of parse and number of phonemes for words in Experiment 2 and Treiman and Chafetz (1987)

Method	Number of phonemes	Nononset parse	Onset-rime parse
Space	3	87.7	90.6
		(1.8)	(1.6)
	5	96.5	95.4
		(0.7)	(0.8)
Asterisks	3	91.4	91.9
		(1.4)	(1.1)
	5	93.5	93.8
		(1.0)	(0.9)
Slashes (Treiman and Chafetz, 1987)	3	85.1	86.5
	5	95.8	95.6

Standard errors in parentheses

For accuracy, there were no differences between stimuli parsed by onset-rime and nononset for both the space method of presentation ($F1(1,52)=0.43$, $MSE = 0.009$, $p < .6$, *F2*(1,34)=0.05, *MSE* = 0.004, p < .9) and the asterisks method of presentation (*F1*(1,63)=.14, *MSE* = 0.01, p < .8, *F2*(1,34)=0.94, *MSE* = 0.001, p < .4). For number of phonemes, the difference in accuracy for 3 and 5 phoneme items was significant by participants (*F1*(1,52)=43.64, *MSE* = 0.006, p < .001), and marginal by items (*F2*(1,34)=3.84, *MSE* = 0.01, p < .06) for the space method of presentation. This main effect was due to higher accuracy in the 5 phoneme condition by an average of 4.6%. The effect for phonemes was also marginal for the asterisks method by participants (*F1*(1,52)=3.99, *MSE* = 0.006, p < .06), but not significant for items (*F2*(1,34)=0.18, *MSE* = 0.013, p < .7). This marginal difference by participants was also due to higher accuracy in the 5 phoneme condition by an average of 2%. In addition, parse and

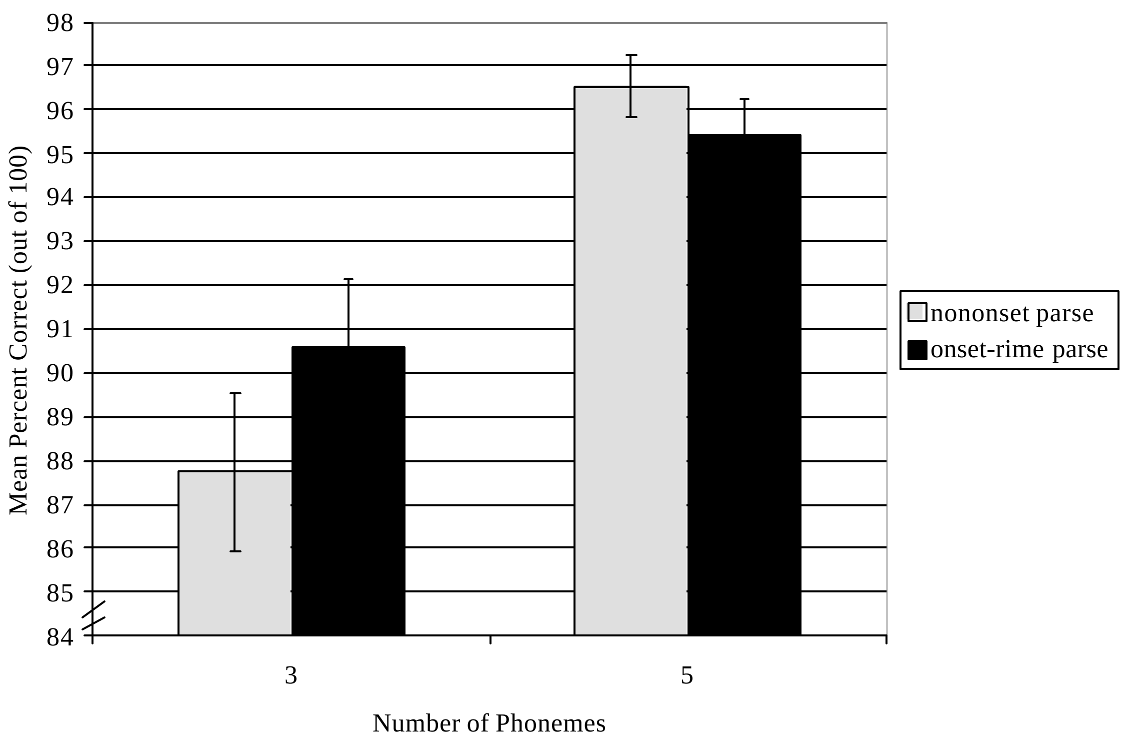


Figure 4.8. Accuracy for type of parse and number of phonemes for words for the space method of presentation.

number of phonemes failed to interact for both the space method of presentation ($F1(1,52)=1.80$, $MSE = 0.01$, $p < .2$, $F2(1,34)=0.74$, $MSE = 0.004$, $p < .4$) and the asterisks method of presentation ($F1(1,63)=.009$, $MSE = 0.01$, $p < 1.0$, $F2(1,34)=0.42$, $MSE = 0.001$, $p < .6$).

For these two methods, the general pattern of results was that 5 phoneme words were recognized with greater accuracy than the three phoneme words. This difference was also not dependent how the item was parsed. One possible reason for this finding is that when the five letters of the stimulus item mapped onto five distinct phonemes, the computation was easier and were better able to make a word response. However if this were the case, reaction times for these items would also be slower than the 5 phoneme

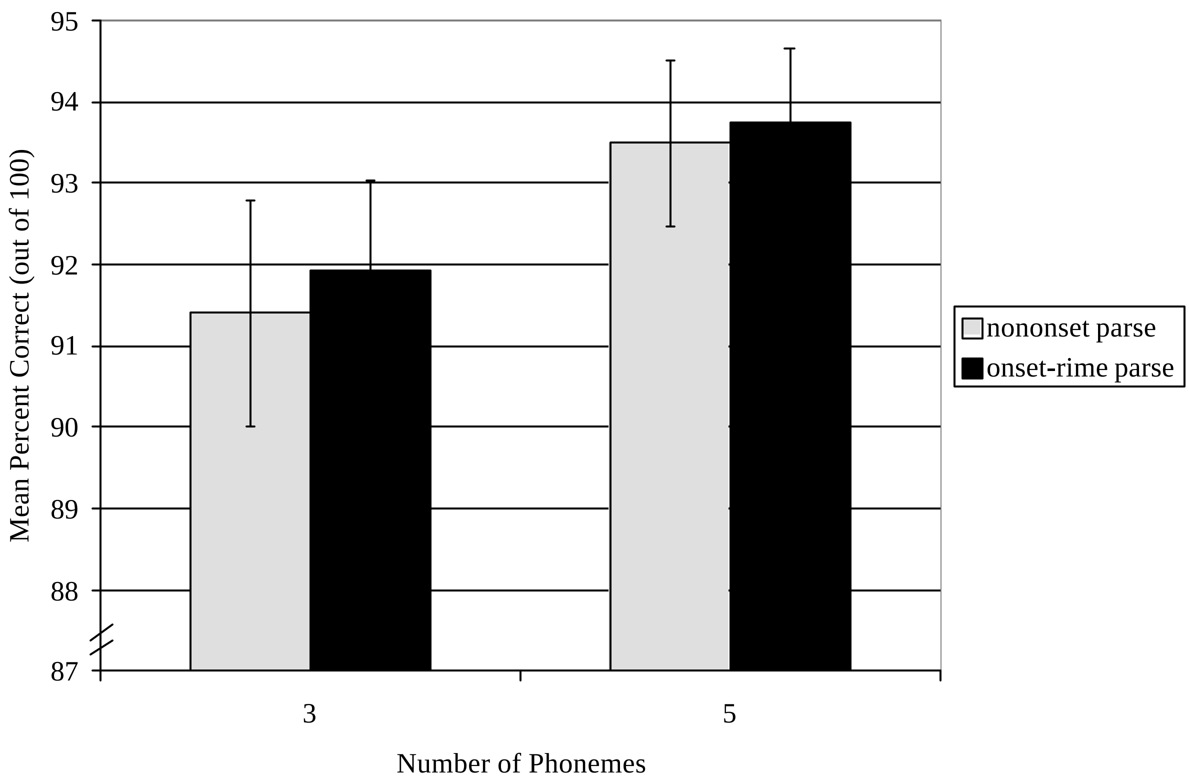


Figure 4.9. Accuracy for type of parse and number of phonemes for words for the asterisks method of presentation.

items. However the reaction time analysis reported earlier found that response times for 3 phoneme words were significantly faster than 5 phoneme words. Therefore the more likely explanation is that this accuracy difference is due to a speed-accuracy tradeoff.

The findings from the accuracy analysis are similar to those reported by Treiman and Chafetz (1987). They also found significant accuracy differences between 3 and 5 phoneme words. In their experiment, accuracy was relatively consistent across type of parse. However for number of phonemes, as was found in the current analysis, participants were more accurate for the 5 phoneme condition than the 3 phoneme condition by an average of 10%.

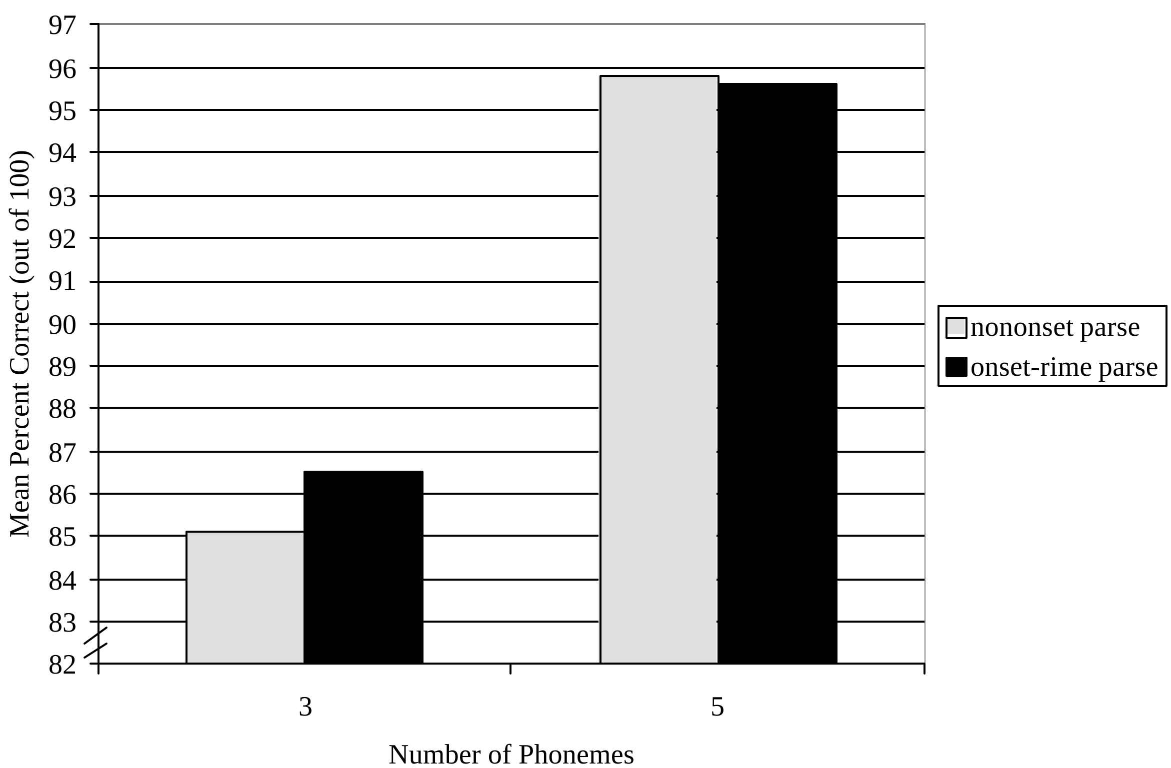


Figure 4.10. Accuracy for words by type of parse and number of phonemes for Treiman and Chafetz (1987).

Although Treiman and Chafetz (1987) failed to obtain significant differences in reaction time for 3 and 5 phoneme words, on average, response times to 5 phoneme words were 57 ms faster than responses to 3 phoneme words. Taking these reaction time results into account it appears that the effects for number of phonemes in both experiments were the result of a speed-accuracy tradeoff.

Accuracy by phoneme and parse for nonwords

The effects for parse and number of phonemes were also examined for nonword accuracy. As in the word accuracy analysis, the data were split by method of presentation. The analysis for the space method of presentation revealed no effects of number of phonemes, ($F1(1,52)=1.12$, $MSE = 0.01$, $p < .3$, $F2(1,34)=0.00$, $MSE = 0.01$, $p < 1.0$), parse ($F1(1,52)=0.37$, $MSE = 0.02$, $p < .6$, $F2(1,34)=0.00$, $MSE = 0.006$, $p < 1.0$),

Table 4.6

Accuracy (out of 100%) as a function of type of parse and number of phonemes for nonwords in Experiment 2 and Treiman and Chafetz (1987)

Method	Number of phonemes	Nononset parse	Onset-rime parse
Space	3	89.6 (1.8)	89.3 (1.8)
	5	88.7 (1.5)	86.9 (1.5)
Asterisks	3	94.5 (1.2)	90.1 (1.7)
	5	85.9 (1.4)	88.5 (1.3)
Slashes (Treiman and Chafetz, 1987)	3	88.7	88.0
	5	94.6	94.6

Standard errors in parentheses

and no interaction of phoneme and parse (*F1*(1,52)=0.19, *MSE* = 0.01, p < .7, *F2*(1,34)=0.16, *MSE* = 0.006, p < .7).

For the asterisks method, the results were quite different. Like the previous analysis, there was no effect of parse by participants (*F1*(1,63)=0.52, *MSE* = 0.01, p < .5), or items (*F2*(1,34)=0.60, *MSE* = 0.002, p < .5). However there was a main effect of number of phonemes by participants, (*F1*(1,63)=14.36, *MSE* = 0.01, p < .001), but not items (*F2*(1,34)=1.86, *MSE* = 0.01, p < .2). This main effect by participants was due to a 5% greater accuracy for 3 phoneme over 5 phoneme words.

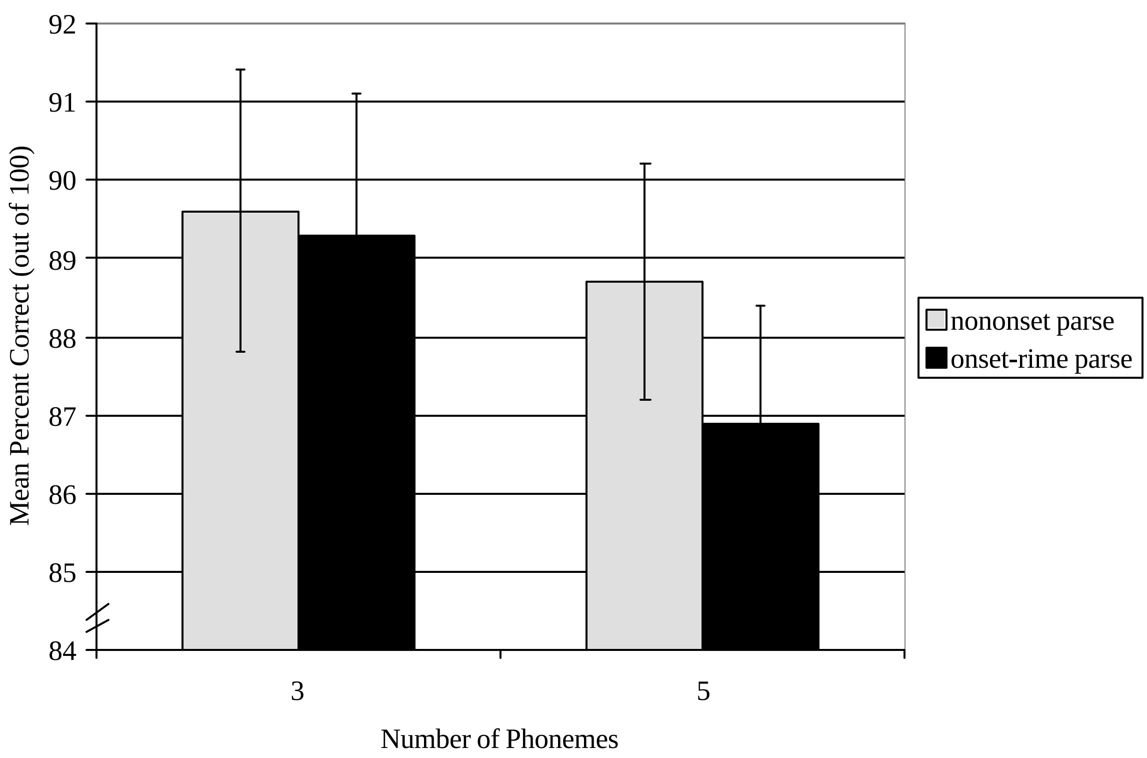


Figure 4.11. Accuracy for nonwords by type of parse and number of phonemes for the space method of presentation.

This main effect was qualified however by an interaction with number of phonemes and parse. The interaction was significant by participants (*F1*(1,63)=5.71, *MSE* = 0.01, p < .02), and but not by items (*F2*(1,34)=2.46, *MSE* = 0.002, p < .2).

The effect for participants was due to significant differences between the 3 and 5 phoneme conditions for the nononset parsed items ($t_{63} = 4.80$, $p < .001$). For nononset parsed items, accuracy for 3 phoneme nonwords was higher than 5 phoneme nonwords by an average of 6.3%. For onset-rime parsed items, the 2.2% difference in the same direction was not significant ($t_{63} = .72$, $p < .5$). It is unclear why this interaction occurred and why it only occurred for the asterisks presentation, however it is possible that the masking provided by the asterisks increased the difficulty of judgment for these items and that the difficulty of the judgment plus the difficulty of mapping five letters onto three phonemes may have made it more likely for participants to identify these items as nonwords. However, these nonwords also had the fastest response times which makes the previous explanation unlikely and a speed accuracy tradeoff a more likely explanation.

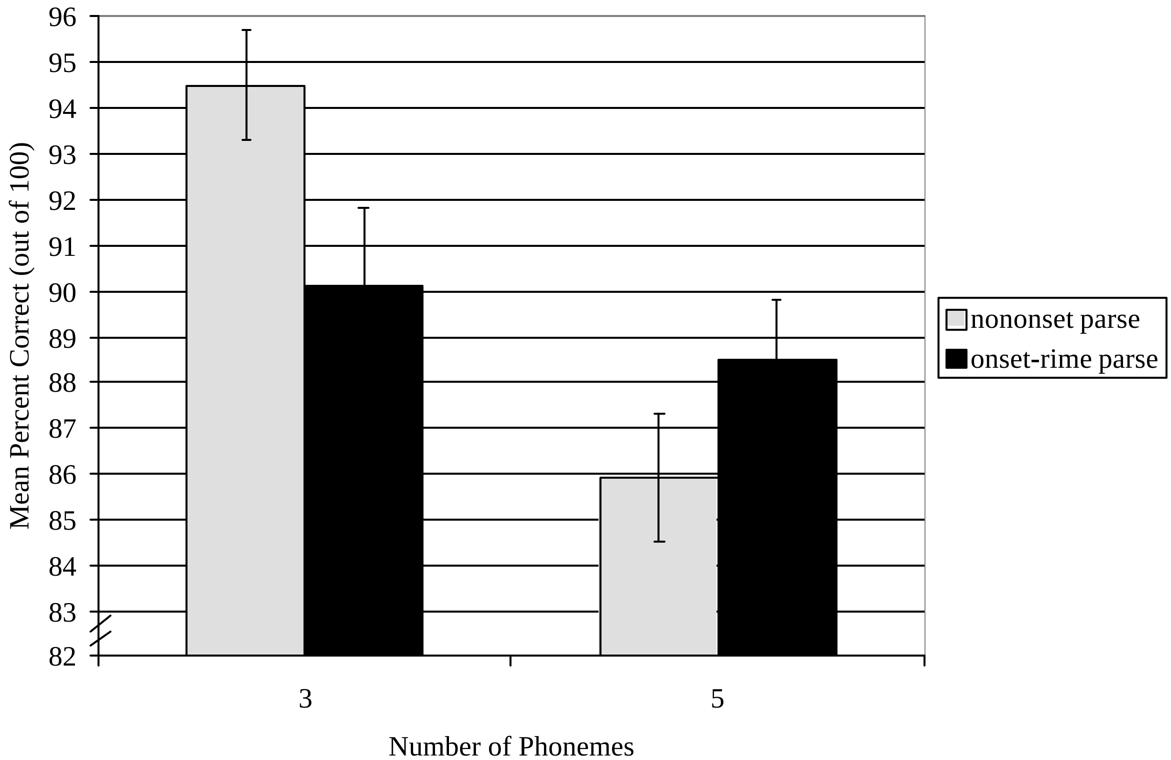


Figure 4.12. Accuracy for nonwords by type of parse and number of phonemes for the asterisks method of presentation.

For nonwords Treiman and Chafetz (1987) also found significant differences between 5 phoneme stimuli and 3 phoneme stimuli. However, in their study, 5 phoneme nonwords were more accurate than three phoneme nonwords. In general for nonwords, the participants in Treiman and Chafetz's study were also slower for their responses to nonwords than the participants in Experiment 2. This pattern again is consistent with the hypothesis that Treiman and Chafetz's (1987) participants may have been performing a check before responding to the stimulus. Therefore the criteria used to make nonword judgments between the two groups may have been different and the cause of the divergent findings with their study and Experiment 2.

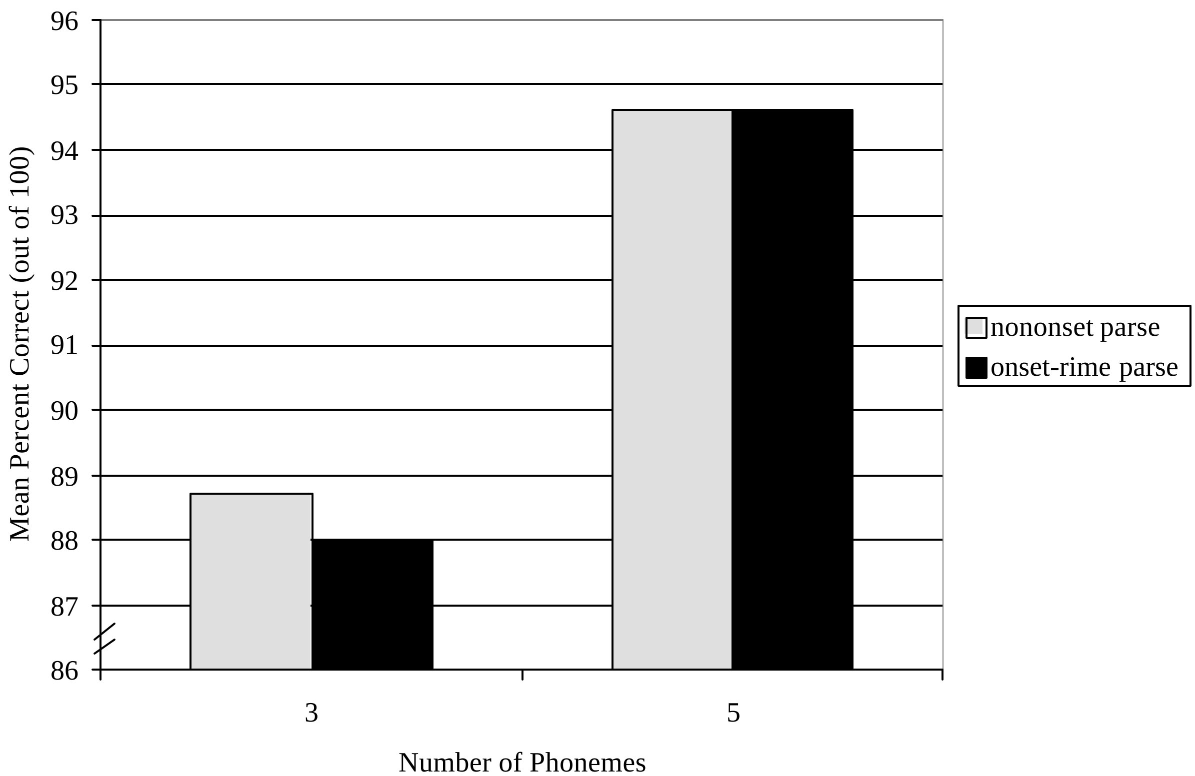


Figure 4.13. Accuracy for nonwords by type of parse and number of phonemes for Treiman and Chafetz (1987).

Accuracy comparison by SOA for words

To determine whether there were any effects of parse and number of phonemes on accuracy over the three SOA conditions, and whether these effects differed by SOA, a 2 (parse) X 2 (number of phonemes) X 3 SOA was conducted. This analysis revealed a main effect of number of phonemes, which was significant by participants ($F1(1,114)=30.55$, $MSE = 0.007$, $p < .001$), but not by items ($F2(1,34)=1.39$, $MSE = 0.03$, $p < .3$). There was also an interaction between parse and SOA that was significant by participants ($F1(2,144)=3.10$, $MSE = 0.008$, $p < .05$), but not by items ($F2(2,68)=0.34$, $MSE = 0.003$, $p < .8$). These two effects however were qualified by a three-way

Table 4.7

Accuracy (out of 100%) as a function of type of parse and number of phonemes for words by SOA in Experiment 2

SOA (ms)	Number of phonemes	Nononset parse	Onset-rime parse
50	3	92.7	87.5
		(2.1)	(1.8)
	5	93.8	94.8
		(1.3)	(1.2)
250	3	86.0	93.0
		(1.8)	(1.5)
	5	95.2	94.2
		(1.1)	(1.0)
500	3	91.3	92.5
		(1.8)	(1.5)
	5	95.4	94.6
		(1.1)	(1.0)

Standard errors in parentheses

interaction of parse, number of phonemes, and SOA. This interaction occurred because the three different SOAs yielded three different patterns of results.

The 50 ms SOA analysis revealed a main effect of number of phonemes by participants ($F1(1,31)=9.92$, $MSE = 0.006$, $p < .02$), but not by items ($F2(1,35)=1.33$, $MSE = 0.008$, $p < .9$). This main effect for participants was due to higher accuracy for 5 phoneme words over 3 phoneme words by an average of 4.2%. There was no main effect for parse by participants ($F1(1,31)=1.94$, $MSE = 0.0071$, $p < .6$), or items ($F2(1,35)=0.02$, $MSE = 0.02$, $p < 1.0$) for this SOA. However the main effect of phonemes was qualified by an interaction with type of parse that was marginal for participants ($F1(1,31)=3.21$, $MSE = 0.01$, $p < .09$), but not by items ($F2(1,35)=0.16$, $MSE = 0.02$, $p < .7$). This marginal interaction was due to greater accuracy for 5 phoneme words compared with 3 phoneme words for onset-rime parsed stimuli ($t_{31} = 3.57$, $p < .002$). On average 5 phoneme words were 7.3% more accurate than three phoneme words for this type of parse.

For 250 ms there was also a main effect for number of phonemes by participants ($F1(1,42)=20.53$, $MSE = 0.006$, $p < .003$), but not by items ($F2(1,35)=1.04$, $MSE = 0.02$, $p < .4$). This main effect for participants was the result of higher accuracy for 5 phoneme words over 3 phoneme words by an average of 5.1%. The effect for parse was not significant by participants, ($F1(1,42)=3.90$, $MSE = 0.01$, $p < .2$), or by items ($F2(1,35)=4.14$, $MSE = 0.002$, $p < .2$). The main effect however were qualified by an interaction between phoneme and parse that was marginal both for participants ($F1(1,42)=5.55$, $MSE = 0.01$, $p < .09$), and items ($F2(1,35)=7.00$, $MSE = 0.002$, $p < .06$). This interaction was marginal because for the onset-rime parsed items, there were no significant accuracy differences between 3 and 5 phoneme words. However for nononset rime parsed items, accuracy for 3 phoneme words was significantly lower than 5 phoneme words by an average of 9.1% ($t_{42} = 3.88$, $p < .001$).

In the 500 ms SOA condition, there was a main effect of number of phonemes which was marginal by participants ($F1(1,41)=5.16$, $MSE = 0.008$, $p < .09$), but not items ($F2(1,35)=1.24$, $MSE = 0.01$, $p < .9$). The marginal effect by participants was due to higher accuracy for 5 phoneme words over 3 phoneme words by an average of 4.1%. For this SOA condition there was no effect of type of parse by participants ($F1(1,41)=0.03$, $MSE = 0.006$, $p < 1.0$), or items ($F2(1,35)=0.12$, $MSE = 0.003$, $p < 1.0$), and these two variables did not interact (*F1*(1,42)=0.41, *MSE* = 0.01, $p < .6$, *F2*(1,35)=1.58, *MSE* = 0.003, $p < .9$).

For these analyses, once again 5 phoneme words were more accurate than 3 phoneme words. This finding was fairly consistent across all three SOAs. For the 250 SOA condition, there was also a marginal interaction of number of phonemes and type of parse. This interaction revealed that at 250 ms, there were no accuracy differences for onset-rime parsed stimuli, the accuracy difference for number of phonemes was only present for nononset parsed stimuli.

Accuracy comparison by SOA for nonwords

The nonword analysis revealed a main effect of number of phonemes that was significant by participants (*F1*(1,114)=11.14, *MSE* = 0.01, $p < .002$), but not by items (*F2*(1,33)=0.54, *MSE* = 0.04, $p < .5$). This main effect by participants was due to greater accuracy for 3 phoneme nonwords over 5 phoneme nonwords by an average of 3.4%. There was also a main effect of SOA that was not significant by participants (*F1*(2,114)=0.02, *MSE* = 0.02, $p < 1.0$), but was significant by items (*F2*(2,68)=428.13, *MSE* = 0.002, $p < .001$). This main effect by participants was due to significant differences between all three SOA conditions. First, responses to nonwords in the 250 ms condition were an average of 23.2% more accurate than responses in the 50 ms condition ($t_{35} = 26.07$, $p < .001$) and an average of 2.4% less accurate than the 500 ms condition ($t_{35} = 3.18$, $p < .01$). In addition, response times to nonwords in the 500 ms SOA condition were significantly more accurate than responses in the 50 ms condition ($t_{35} = 25.11$, $p < .001$).

Table 4.8

Accuracy (out of 100%) as a function of type of parse and number of phonemes for nonwords by SOA in Experiment 2

SOA (ms)	Number of phonemes	Nononset parse	Onset-rime parse
50	3	90.1	90.6
		(2.1)	(2.4)
	5	87.8	87.8
		(1.9)	(1.9)
250	3	95.0	89.5
		(1.8)	(2.0)
	5	87.2	85.5
		(1.7)	(1.6)
500	3	91.3	89.3
		(1.8)	(2.1)
	5	86.7	90.3
		(1.7)	(1.6)

Standard errors in parentheses

The findings from the nonword analysis by SOA converge with the results of the word analysis by SOA. In the word analysis in general, 5 phoneme words had higher response accuracy than 3 phoneme words. In the nonword analysis, this finding is reversed, and 3 phoneme nonwords have higher accuracy than 5 phoneme words. For type of parse, in the word analysis, there was some evidence that onset-rime parsed stimuli were processed more accurately than nononset parsed items. However for the nonword analyses, when the effect was present, it was in the opposite direction.

Although the reaction time data did not provide much evidence that onset-rime parsed words were processed more easily than nononset parsed words, the accuracy data does provide some evidence that ultimate decision for these words was more accurate than for those parsed after the first vowel. This finding for accuracy does provide some evidence that processing was easier or ultimately more successful when parsed by onset and rime and that this parsing may help guide word recognition for native speakers of English. However the question that remains is whether the German-English bilinguals also show similar effects, or whether this group does not process via the same sublexical units as native speakers of English.

Summary of Native English results

Overall the findings across both methods of presentation and SOAs diverge from the results reported by Treiman and Chafetz (1987). Native English participants from Experiment 2 did not show the predicted benefit for onset-rime parsed words and nonwords. Instead, the benefit was in the opposite direction, nononset parsed stimuli were processed faster than onset-rime parsed stimuli. Although the participants in this study demonstrated the opposite effect from that reported by Treiman and Chafetz, the presence of a difference between these two conditions suggests that there is a critical parsing difference between the nononset parse and onset-rime parse conditions in Experiment 2 for native English speakers.

Comparing the RT results from these two experiments also suggests that different task strategies may have caused the divergent findings across these two experiments. The slower reaction times for the participants in Treiman and Chafetz's experiment point to a strategy that was occurring relatively late in processing where participants were checking their responses to determine whether or not the slashed string of letters was a word or a nonword. Because the participants in this experiment had access to the all of the letters of the string at once, having the slashed presentation may have made the rime more salient and a strong indicator of an item's lexicality.

Although the data from the native English speakers from Experiment 2 fail to replicate those of Treiman and Chafetz (1987), the question of whether or not the native German participants in this experiment will demonstrate similar effects to the native English participants is still an interesting question. If native German speakers do demonstrate effects similar to those of native English speakers, it suggests that they are not relying on L1 processing strategies to complete this task.

Native German Speaker Group

Exclusion of items and replacement of missing condition means

Prior to conducting the analyses, one stimulus item, GNASH, was excluded from the data. This item was excluded because of the 36 participants remaining in this task, only two participants responded to this item with a correct "yes" response. In addition to the removal of this item, in the item matrix several condition means needed to be replaced due to the lack of a correct response for a particular item in one condition. This is perhaps not completely surprising because each of the 36 remaining participants completed one of eight scripts. This division left five participants for six of the eight scripts and three participants for the remaining two scripts. Of the 576 item means for this group (72 stimulus items X 8 scripts), seven means were replaced for six items that included both words and nonwords (BLIMP (2), BRUST, CRANK, CRIST, FLUNK, and SHACK). To avoid reducing the variability within any particular condition, these means were replaced by taking the average of the existing condition means within each item.

Reaction time for words by phoneme and parse

As in the native English group analysis, the data from the native German group was first examined by method of presentation. The data were then analyzed using a 2 (type of parse) X 2 (number of phonemes) ANOVA. The space method of presentation revealed a marginal effect of type of parse by participants ($F1(1,15)=4.05$, MSE = 2677.6, $p < .07$), but not items ($F2(1,33)=1.51$, MSE = 6346.0, $p < .3$). This marginal effect was due to faster response times for nononset parsed words over onset parsed words by 26 ms on average.

Table 4.9

Reaction Time (ms) for native German speakers as a function of type of parse and number of phonemes for words in Experiment 2

Method	Number of phonemes	Nononset parse	Onset-rime Parse	Parse difference
Space	3	699	716	-17
		(35.1)	(24.3)	
	5	726	761	-35
		(35.1)	(34.1)	
Asterisks	3	763	786	-23
		(35.1)	(34.3)	
	5	855	894	-39
		(49.7)	(51.1)	

Standard errors in parentheses

This analysis also revealed a main effect of the number of phonemes that was significant by participants ($F1(1,15)=5.83$, MSE = 3676.2, $p < .03$), but only marginal by items ($F2(1,33)=3.41$, MSE = 11951.5, $p < .08$). As in the native English analysis, this main effect reflected significantly faster response times to 3 phoneme words compared to

5 phoneme words. The size of this difference was 56 ms. These two variables did not interact either for participants (*F1*(1,15)=0.22, *MSE* = 5633.5, p < .7), or items (*F2*(1,33)=.15, *MSE* = 6346.0, p < .8).

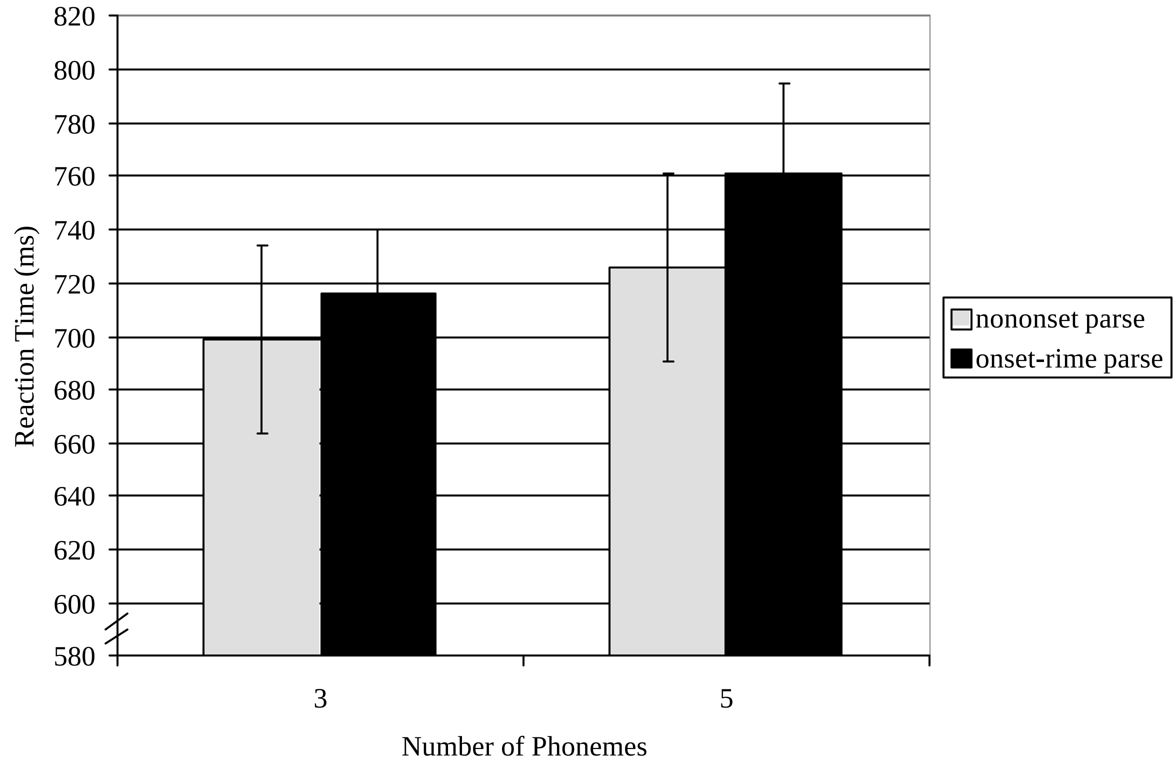


Figure 4.14. Word reaction time (ms) for native German speakers by type of parse and number of phonemes for the space method of presentation.

For the asterisks method of presentation, the pattern of results was quite similar. This analysis revealed an effect of parse that was marginal by participants (*F1*(1,19)=3.75, *MSE* = 5098.1, p < .07), but not items (*F2*(1,33)=1.70, *MSE* = 23519.0, p < .3). As for the space method of presentation, this marginal effect was due to an average difference of 47 ms between nononset parsed words and onset-rime parsed words. Once again, onset-rime parsed words had slower response times than nononset parsed words. In addition, there was also a main effect of number of phonemes that was significant both by participants (*F1*(1,19)=7.53, *MSE* = 26523.0, p < .02), and items (*F2*(1,33)=8.98, *MSE* = 24987.1, p < .006). This main effect was due to faster response times for 3 phoneme

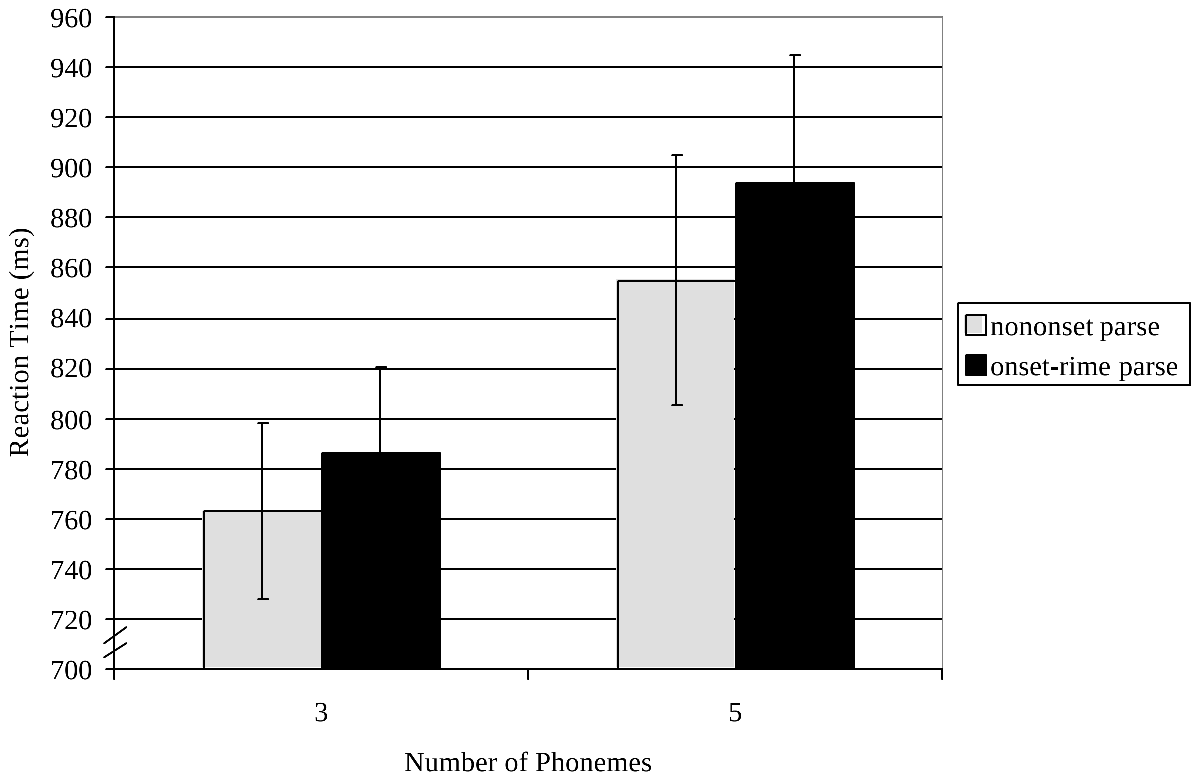


Figure 4.15. Word reaction time (ms) for native German speakers by type of parse and number of phonemes for the space method of presentation.

words over 5 phoneme words by an average of 100 ms. Once again, for this analysis parse and number of phonemes failed to interact by participants (*F1*(1,19)=0.33, *MSE* = 3816.2, p < .6), and items (*F2*(1,33)=0.27, *MSE* = 23519.0, p < .7).

The results from these two methods of presentation have a high degree of overlap both with each other and with the results from the native English speakers. Both groups demonstrated slower response times for 5 phoneme words than 3 phoneme words. There was also some evidence that processing was influenced by parse although this evidence was marginal. For both native English and German speakers, the parsing effects were always in the opposite direction to those reported by Treiman and Chafetz (1987). Although the pattern of results is the same, one difference for these groups is that the effect size for the native German speakers was much larger for number of phonemes than the native English speaker group. In this case, the difference for the native German group was twice as large. Evidence for a greater reliance on smaller unit processing strategies is somewhat supported by the greater difference between 3 and 5 phoneme words for the native German group. Processing at a smaller grain size should lead to larger effects between these two conditions, and the results from these two sets of analyses show that the effect for native German speakers is roughly double the effect shown by native English speakers.

Although it is impossible from the results to rule out completely that the native German group was able mostly to use L2 processing strategies, it is not surprising that L1 strategies may have been more active during this lexical decision task because the parsing of the words made the task more demanding cognitively than a standard lexical decision task. This extra cognitive load may have increased the probability that the native German participants would be somewhat affected by word length and thus show greater benefits for the 3 phoneme word condition over the 5 phoneme condition.

Reaction time comparison by SOA for words

To examine whether the timing of processing imposed by the experiment affected lexical parsing in this group, the effects of number of phonemes and parse were also examined including SOA as a between-participants variable in the participant analysis and as a within-item variable in the item analysis. For this native speaker group, only two levels of the SOA were used: 50 ms and 500 ms.

Table 4.10

Reaction time (ms) for native German speakers as a function of type of parse and number of phonemes for words by SOA in Experiment 2

SOA (ms)	Number of phonemes	Nononset parse	Onset-rime parse
50	3	732	762
		(36.2)	(32.1)
	5	852	865
		(45.7)	(47.8)
500	3	738	748
		(36.2)	(32.1)
	5	745	806
		(45.7)	(47.8)

Standard errors in parentheses

The analysis by SOA revealed a main effect of parse that was significant by participants (*F1*(1,34)=7.47, *MSE* = 3983.6, p < .01), but only marginal by items (*F2*(1,33)=3.48, *MSE* = 12752.4, p < .08). Like the native English group, this main effect of parse was due to faster response times to nononset parsed words by an average of 29 ms. There was also a main effect of number of phonemes that was significant by participants (*F1*(1,34)=11.72, *MSE* = 15833.7, p < .002), and items (*F2*(1,33)=8.02, *MSE* = 28459.0, p < .008). This main effect for number of phonemes was due to faster response times for 3 phoneme words than for 5 phoneme words.

There was also an effect of SOA that was not significant by participants (*F1*(1,34)=0.69, *MSE* = 96677.4, p < .5), but was significant by items (*F2*(1,33)=10.90, *MSE* = 6677.0, $p < .003$). This main effect by items was due to slower response times in the 50 ms condition than in the 500 ms condition. Finally these significant main effects were qualified by an interaction between number of phonemes and SOA that was marginal by participants ($F1(1,17)=8.73$, *MSE* = 15833.7, $p < .07$), and significant by items ($F2(1,33)=8.41$, *MSE* = 6677.0, $p < .007$). This interaction by items was significant because 3 phoneme words were only faster than 5 phoneme words for the 50 ms SOA ($t_{33} = 3.59$, $p < .003$).

The main effects of parse and number of phonemes for the native German speaker group resemble the word reaction time data for native speakers of English. The one main difference is that the effect for number of phonemes changed over the SOA conditions for the German group but not the native English speaker group. For this group response times to 5 phoneme words in the 50 ms condition disrupted processing more than in any other condition. This was evidenced by a 130 ms difference in response time for 3 phoneme words and 5 phoneme words at the 50 ms SOA. For the 500 ms SOA, the mean difference between 3 and 5 phoneme words was 32 ms, a difference that was not significant. This interaction suggests that these participants may have been somewhat reliant on L1 processing strategies and that these strategies may have been unfolding over a different time course than the native English speakers.

Reaction time comparison by method for nonwords

For nonwords, performance across type of parse and number of phonemes was highly consistent. Nonword data was analyzed by a 2 (parse) X 2 (number of phonemes) ANOVA for both the space method of presentation and the asterisks method of presentation. There was no effect of type of parse for either the space method of presentation ($F1(1,15)=1.74$, $MSE = 1815.1$, $p < .3$, $F2(1,34)=0.08$, $MSE = 5029.8$, $p < .8$), or the asterisks method of presentation ($F1(1,19)=0.54$, $MSE = 9852.2$, $p < .5$, $F2(1,34)=0.07$, $MSE = 15787.6$, $p < .8$).

Table 4.11

Reaction Time (ms) for native German speakers as a function of type of parse and number of phonemes for nonwords in Experiment 2

Method	Number of phonemes	Nononset parse	Onset-rime parse
Space	3	795	812
		(40.5)	(51.5)
	5	815	827
		(48.3)	(43.9)
Asterisks	3	972	991
		(65.1)	(73.3)
	5	996	1010
		(69.3)	(64.8)

Standard errors in parentheses

These analyses also failed to find an effect of number of phonemes in the space method ($F1(1,15)=0.44$, $MSE = 11197.7$, $p < .6$, $F2(1,34)=0.13$, $MSE = 7175.3$, $p < .8$), or the asterisks method ($F1(1,19)=0.44$, $MSE = 21328.0$, $p < .6$, $F2(1,34)=0.11$, $MSE = 16576.6$, $p < .8$). These two variables also did not produce a significant interaction for the space analysis (*F1*(1,15)=0.04, *MSE* = 2812.6, p < .9, *F2*(1,34)=0.03, *MSE* = 5029.9, p < .9), or the asterisks analysis (*F1*(1,19)=0.03, *MSE* = 5430.5, p < .9, *F2*(1,34)=0.00, *MSE* = 15787.6, p < 1.0).

Reaction time comparison by SOA for nonwords

For SOA, the findings were generally similar. The 2 (parse) X 2 (number of phonemes) X 2 (SOA) revealed no effect of parse (*F1*(1,34)=1.42, *MSE* = 5933.7, p < .3, *F2*(1,34)=0.00, *MSE* = 1368.2, p < .7) and no effect of number of phonemes (*F1*(1,34)=0.87, *MSE* = 16433.5, p < .4, *F2*(1,34)=0.21, *MSE* = 12744.4, p < .7). There was a main effect of SOA that was not significant by participants (*F1*(1,34)=0.27, *MSE* = 272799.2, p < .7), but was significant by items (*F2*(1,34)=4.80, *MSE* = 10408.8, p < .4). This main effect by items was due to faster response times to nonwords in the 500 ms SOA condition by and average of 39 ms.

Table 4.12

Reaction time (ms) for native German speakers as a function of type of parse and number of phonemes for nonwords by SOA in Experiment 2

SOA (ms)	Number of phonemes	Nononset parse	Onset-rime parse
50	3	907	922
		(61.0)	(69.5)
	5	966	943
		(65.0)	(62.0)
500	3	880	900
		(61.0)	(69.5)
	5	866	914
		(65.0)	(61.9)

Standard errors in parentheses

Findings from the native German group in this analysis diverge somewhat from those findings from the native English speaker group. Unlike the native English speaker group, there were no effects of number of phonemes in the nonword reaction time analysis and no effects of type of parse. Together with word analyses, these findings suggest that the native German speaker group was using processing strategies that were different than those used by the native English speakers.

Accuracy comparison by method for words

Accuracy for words on this task was examined using a 2 (parse) by 2 (number of phonemes) ANOVA. Separate analyses were conducted for the space and asterisks methods of presentation. For the space method of presentation, the analysis revealed a main effect of number of phonemes that was significant by participants, ($F1(1,15)=18.30$, $MSE = 0.01$, $p < .001$), but not significant by items $F2(1,33)=2.71$, $MSE = .09$, $p < .2$). There was no effect of parse by participants ($F1(1,15)=0.39$, $MSE = 0.02$, $p < .6$), or by items $F2(1,33)=0.008$, $MSE = .02$, $p < 1.0$). However there was a marginal interaction between parse and phoneme for participants ($F1(1,15)=3.38$, $MSE = 0.008$, $p < .09$), but not by items $F2(1,33)=0.59$, $MSE = .02$, $p < .5$).

For the asterisks method of presentation, there was also a significant main effect of number of phonemes in the participant analysis ($F1(1,19)=28.24$, $MSE = 0.006$, $p < .001$), but not in the item analysis ($F2(1,33)=1.65$, $MSE = .09$, $p < .3$). There was no effect of parse by participants ($F1(1,19)=0.75$, $MSE = 0.02$, $p < .4$), or by items $F2(1,33)=1.10$, $MSE = .007$, $p < .4$); and parse and number of phonemes did not interact ($F1(1,15)=1.10$, $MSE = 0.01$, $p < .4$, $F2(1,33)=1.10$, $MSE = .007$, $p < .4$).

The main effect for number of phonemes for both of these tasks was the result of higher accuracy in the 3 phoneme condition over the 5 phoneme condition. For the space method of presentation there was an 11.1% average difference between these two conditions, and for the asterisks method of presentation, the average difference was 9.5%.

Table 4.13

Accuracy (out of 100%) for native German speakers as a function of type of parse and number of phonemes for words in Experiment 2

Method	Number of phonemes	Nononset parse	Onset-rime parse
Space	3	86.3	92.3
		(3.5)	(2.7)
	5	79.2	77.1
		(3.0)	(2.9)
Asterisks	3	93.2	88.3
		(2.6)	(3.0)
	5	81.3	81.3
		(2.3)	(2.8)

Standard errors in parentheses

The marginal interaction for the space method of presentation seems to have been due to this difference between the 3 phoneme condition and the 5 phoneme condition for onset-rime parsed words only being reliable for onset-rime parsed words.

Accuracy comparison by SOA for words

The results for SOA were quite similar to the results split by method. For this analysis there were significant differences for number of phonemes by participants ($F1(1,34)=45.45$, $MSE = 0.008$, $p < .001$), but not by items $F2(1,33)=2.27$, $MSE = .17$, p

< .2).This main effect was due to a 10.2% average greater accuracy for the 3 phoneme condition over the 5 phoneme condition. Aside from this main effect, no other effects or interactions were significant.

Table 4.14

Accuracy (out of 100%) for native German speakers as a function of type of parse and number of phonemes for words by SOA in Experiment 2

SOA (ms)	Number of phonemes	Nononset parse	Onset-rime parse
50	3	92.4	88.7
		(3.1)	(2.9)
	5	80.6	77.8
		(2.7)	(2.9)
500	3	87.8	91.5
		(3.1)	(2.9)
	5	80.1	81.0
		(2.7)	(2.9)

Standard errors in parentheses

These results for word accuracy across both of these methods are at odds with the findings from Treiman and Chafetz (1987) and the native English speakers in the present experiment. For the native German speaker group, accuracy was greater for the 3 phoneme condition than the 5 phoneme condition. This difference may have occurred for a number of reasons. First, these results may be revealing different processing preferences for the native German speaker group. From the reaction time data, it is clear that not only were 3 phoneme words more accurate, but participants also responded to these words faster than words with 5 phonemes. Therefore it is possible that these items were easier to process. However, as mentioned earlier, the words in the three phoneme condition were more accurate than the words in the 5 phoneme condition so this difference may be due to greater familiarity for these items as nonnative speakers of English.

Accuracy comparison by method for nonwords

For nonwords the two methods of presentation yielded different results. For the space method of presentation, there were no effects of number of phonemes, ($F1(1,15)=2.40$, $MSE = 0.01$, $p < .2$, $F2(1,34)=0.80$, $MSE = .05$, $p < .4$), type of parse, ($F1(1,15)=0.06$, $MSE = 0.02$, $p < .9$, $F2(1,34)=0.005$, $MSE = .01$, $p < 1.0$), or interaction between number of phonemes and parse ($F1(1,15)=0.77$, $MSE = 0.02$, $p < .4$, $F2(1,34)=0.76$, $MSE = .01$, $p < .4$).

Table 4.15

Accuracy (out of 100%) for native German speakers as a function of type of parse and number of phonemes for nonwords in Experiment 2

Method	Number of phonemes	Nononset parse	Onset-rime parse
Space	3	90.6	88.5
		(2.6)	(4.7)
	5	83.3	87.0
		(2.3)	(3.1)
Asterisks	3	93.3	89.2
		(2.2)	(2.5)
	5	82.1	77.9
		(2.4)	(2.7)

Standard errors in parentheses

For the asterisks method of presentation however there was a main effect of number of phonemes that was significant for participants (*F1*(1,19)=21.61, *MSE* = 0.01, p < .001), but marginal by items (*F2*(1,34)=3.44, *MSE* = .06, p < .08). For this method of presentation, 3 phoneme nonwords were more accurate than 5 phoneme nonwords by an average of 11.2%. As in the space method analysis, this method also found no effects of parse (*F1*(1,19)=2.18, *MSE* = 0.02, p < .2, *F2*(1,34)=2.40, *MSE* = .01, p < .2), and no interaction between phoneme and parse (*F1*(1,19)=0.00, *MSE* = 0.01, p < 1.0, *F2*(1,34)=0.00, *MSE* = .01, p < 1.0).

Accuracy comparison by SOA for nonwords

The results for nonword accuracy across SOAs follow closely what was found across method of presentation. First, there was an effect of number of phonemes that was significant by participants (*F1*(1,34)=18.84, *MSE* = 0.01, p < 0.001), but not by items (*F2*(1,34)=2.09, *MSE* = .10, p < 0.2). As was the case for the word accuracy analysis, for these analyses, 3 phoneme words were on average 8.2% more accurate than 5 phoneme words. No other effects or interactions approached significance for this analysis.

For both words and nonwords, the consistent finding for the native German speaker group is that 3 phoneme words and nonwords are both more accurate and faster than 5 phoneme words and nonwords. For reaction time for words, the native German speakers showed a similar effect for parse as the native English speakers, although the difference did not reach statistical significance for the German group. However the overlap in findings between these two groups for parse was not consistent over the other analyses. It is possible that this effect did not emerge because German native speakers were using

Table 4.16

<u>Accuracy (out of 100%) for native German speakers as a function of type of parse and number of phonemes for nonwords by SOA in Experiment 2</u>

SOA (ms)	Number of phonemes	Nononset parse	Onset-rime parse
50	3	92.6 (2.4)	86.1 (3.5)
	5	85.2 (2.3)	81.9 (3.1)
500	3	91.7 (2.4)	91.7 (3.5)
	5	80.1 (2.3)	81.9 (3.1)

Standard errors in parentheses

processing strategies that were different than the native English group. However because this analysis lacks the 250 ms SOA condition, it is difficult to determine whether this is the case, or whether the window for this effect was not within any of the SOA conditions chosen for the native German speakers.

Summary of native German results

Although Experiment 1 showed a high degree of overlap between the native English speaker and native German speaker groups, the cluster lexical decision task utilized here seemed to distinguish some differences. For this task, it is possible that native German speakers were relying more on smaller unit processing strategies from German, from which emerged a pattern of results that was distinct from the native English speakers.

There are several reasons why native speakers of German may have been somewhat more reliant on strategies in their L1 for this task. First, the unlike the stimulus items from Experiment 1, the items from this task were designed for adult readers. Because these words were lower frequency, they were most likely more difficult for the nonnative speakers to process. In addition, although participants were making responses to complete strings of letters, the clustered presentation plus the SOA between the first cluster and the remainder of the word made this task much more difficult than the standard lexical decision task used in Experiment 1. The increased difficulty of this task seems likely to have increased the cognitive load for the non-native participants in comparison with Experiment 1. Increased cognitive load may also have increased the tendency to rely on L1 strategies.

However, to determine whether these groups did in fact differ, both groups were included in one analysis examining the effects of number of phonemes and parse including method as a between participants variable, and a second including SOA as a between participants variable.

Direct comparison of native English and German groups by method

In this analysis the effects of method, parse, and number of phonemes were compared along with native language group by participants. Because no native German speakers participated in the 250 ms SOA condition, only native speakers of English from the 50 and 500 ms SOA conditions were included in this analysis (n=75). For reaction time, there was an overall main effect of group ($F1(1,107)=15.73$, $MSE = 62821.2$, $p < 0.001$). The native English speakers were faster than the native German speakers by an average of 101 ms. Group however did not interact with any of the other variables in the study suggesting that the additional main effects reported below were similar for both the native English and native German speakers. As in other analyses, this analysis revealed main effects of both phoneme (*F1*(1,107)=32.12, *MSE* = 8910.6, $p < 0.001$), and parse (*F1*(1,107)=18.57, *MSE* = 4371.1, $p < 0.001$).The main effect for number of phonemes was due to faster response times for 3 phoneme words over 5 phoneme words by an average of 55 ms, and for parse, on average nononset parsed items had faster response times than onset-rime parsed items by an average of 29 ms.

In addition to these main effects that were also found in analyses reported earlier, this analysis revealed a main effect of method (*F1*(1,34)=6.08, *MSE* = 62822.0, $p < 0.02$). This main effect was due to overall faster response times for the space method of presentation over the asterisks method of presentation by an average of 63 ms. Although presenting the asterisks after the first onset cluster provided a strong cue as to how many more letters would follow, it seems that these asterisks may have forward masked the remainder of the word, making it more difficult to process the final stimulus item than in the space method of presentation. Finally there was a marginal interaction between method and number of phonemes (*F1*(1,107)=2.04, *MSE* = 8910.6, $p < 0.08$). This interaction was most likely marginal because the difference between 3 phoneme and 5 phoneme words was larger in the asterisks method of presentation than in the space method of presentation (72 vs. 37 ms respectively).The analysis by words suggests that although processing was somewhat slower overall for the native German speakers, the pattern of effects for both speaker groups did not differ from each other.

The same conclusion was reached for the analysis by nonwords. Again, response times were slower for the native German group than the native English group by an average of 49 ms (*F1*(1,107)=5.81, *MSE* = 266962.8, $p < 0.02$). However once again group did not interact with any other variables. In the nonword analysis, only one other effect reached significance, the effect of number of phonemes (*F1*(1,107)=6.54, *MSE* = 13001.2, p < 0.01). As in the word analysis, this effect was due to a 29 ms processing advantage for 3 phoneme words over 5 phoneme words.

As in Experiment 1, the results for reaction time in Experiment 2 did not reveal major differences in processing for the native English and native German speakers. For words, both groups showed faster response times for nononset parsed words, and for both words and nonwords, 3 phoneme stimuli had faster response times than 5 phoneme stimuli. These findings suggest that both of these native speaker groups were using fundamentally similar processing strategies, although the native English group was processing in their L1 and the native German group in their L2.

For accuracy, there were some differences between the two groups however. In addition to an overall difference between the two groups in accuracy (*F1*(1,107)=37.17, *MSE* = 0.01, p < 0.001), there was a significant interaction between number of phonemes and group (*F1*(1,107)=92.33, *MSE* = 0.006, p < 0.001), and a marginal interaction between phoneme, method, and group (*F1*(1,107)=3.24, *MSE* = 0.006, p < 0.08). This interaction between number of phonemes and group was due to a reverse pattern of accuracy for number of phonemes for the two language groups. For both the native English speakers and the native German speakers, there were significant differences between 3 phoneme words and 5 phoneme words. However, for the native English speakers, 5 phoneme words were more accurate by an average of 4.9% ($t_{74} = 5.72$, $p < .001$), and for native German speakers 3 phoneme words were more accurate by an average of 10.2% ($t_{35} = 6.78$, $p < .001$). This accuracy reversal for the two groups may suggest that these two groups were using different processing strategies for ultimately determining whether or not an item was a word. However another more likely explanation is that because the 3 phoneme words were more frequent than the 5 phoneme words, these words may have been more familiar to nonnative speakers.

Finally, for the nonword accuracy analysis, there was a four-way interaction between number of phonemes, parse, method of presentation, and language group, ($F1(1,107)=3.94$, $MSE = 0.01$, $p < 0.05$). However, when the data were analyzed by method for the posthoc comparisons, no effects were significant, and language group did not interact with any variables.

For method of presentation, the one analysis that revealed differences between these two groups was the analysis for word accuracy. The native English group and the native German group showed directly opposite patterns of results. For native English speakers 5 phoneme words were more accurate than 3 phoneme words and for native German speakers, 3 phonemes words were more accurate than 5 phoneme words. Although this difference suggests that the native German group may have been using a different set of accuracy criteria than the native English group, the failure to find differences in the nonword condition suggests that the accuracy difference in the word condition may have been due to greater familiarity for the 3 phoneme items for the native German speakers.

Direct comparison of native English and German groups by SOA

Analyses were also conducted for number of phonemes, parse, and language group with SOA as a between-participants factor for word and nonword reaction time and accuracy. For these analyses, group and SOA did not interact, demonstrating that for the 50 and 500 ms SOA, overall processing looked similar across both groups.

General Discussion

The results of Experiment 3 revealed some minor differences between native English speakers and native German speakers. However in general, the word recognition performance of these two groups was more similar than different. Although reaction time was slower for the native German group, the overall pattern of results did not diverge significantly from the native English speakers. The one somewhat consistent way that these two groups did differ seemed to be for word accuracy. However this difference is the result of differences in English vocabulary for the L1 vs. the L2.

Perhaps the most surprising result from this study was that native English speakers did not show a benefit for having an onset-rime parse presentation. Unlike Treiman and Chafetz's (1987) slash method, neither the space or asterisks methods of presentation showed a benefit for words parsed by onset followed by rime. The presence of this reverse effect in Experiment 2 may have occurred for a number of reasons. First, in the nononset parse condition, during the presentation of the first letter cluster, participants process the onset and then use the vowel as a pointer to possible rime units. When the remaining letters of the word appear, a number of rime candidates have already been activated and then one is ultimately selected. For the time course of processing, 50 ms is too fast to show a large benefit for having this extra letter, but by 250 ms processing has more of a chance to unfold. This advantage would not have occurred in Treiman and Chafetz's experiment, because for the slashed presentation, participants always had access to the entire letter string.

A second possibility is that there were benefits of onset-rime processing in this task. However these benefits were masked by the benefit of having an extra piece of information, i.e., the vowel in the nononset parse case. Because models of word recognition such as the DRC model and BIA+ operate over a series of cycles, and neither model makes assumptions about how long cycles of operation take to complete, it is difficult to determine just how much of a benefit having this extra piece of information would provide for reaction time.

From the data presented here, it is difficult to determine which of these possibilities is most likely, and also what type of lexical information is being activated when participants are being presented with the onset plus the first vowel in the nononset parse condition. Using these same methods of presentation and SOA conditions in a naming experiment and examining the types of vowel errors that people may make throughout the course of the task may provide one way to answer this question. Since word naming will be examined in Experiment 3, the results of this study may shed light on how naming may affect this task. We return to consider this point in the final discussion.

In addition, it is possible that only comparing the 50 and 500 ms SOA conditions for the native English and native German participants did not provide a sensitive enough analysis for identifying differences in the time course of processing for these two groups. Although the native German group in general showed similar patterns in processing across the two methods of presentation and the two SOAs, including the 250 ms condition may help to distinguish processing strategies between these two language groups. Overall, the development of this cluster lexical decision task is still at an early stage. Perhaps with further development, this task may be an effective method for examining differences in lexical parsing and across language groups, and for also identifying the time course of processing within groups and differences in this time course across groups.

In general the findings from Experiment 2 support a bilingual dual-route model more than the BIA+ model. The results from Treiman and Chafetz (1987) that demonstrated a benefit for onset-rime parsed words was probably due to a checking strategy that made the rime more salient. This checking strategy places the locus of the onset-rime benefit late in the processing stream and most likely outside of the lexicon. In contrast, the native English and native German participants from this study did not show this benefit but did show an advantage for nononset parsed items over onset-rime parsed items, suggesting that parsing did occur. Because participants had overall shorter response latencies than the participants from Treiman and Chafetz's study, and because they consistently showed a different pattern of effects, it is most likely that the parsing revealed in Experiment 2 was occurring within the lexicon and possibly within a set of grapheme-to-phoneme correspondence rules. In addition, Experiment 2 revealed differences between 3 and 5 phoneme items. These length differences also support a model like the DRC that has a serial component. This point will be further discussed in the sixth and final chapter.

Chapter 5: Word Length and Body Neighbor Effects in Naming

In Experiments 1 and 2, there were surprisingly few differences between performance for native speakers of English and native speakers of German using English as their L2. Although these two experiments used different manipulations and methods of stimulus presentation, ultimately participants made lexical decisions in each experiment and it is possible that lexical decision is not sensitive enough to distinguish processing differences between native and proficient L2 speakers. For this reason, the third and final experiment compares the same two groups on a word naming task.

Word naming may be more sensitive than lexical decision for revealing parsing preferences because unlike lexical decision, in naming the phonology has to be fully specified and then produced. Because cross-linguistic studies have suggested that native language performance differences between English and German speakers are due to differences in consistency in spelling-to-sound mappings in the two languages, having participants name word and nonword stimuli may provide a clearer picture of how proficient L2 English processing compares to that of native speakers.

The naming task in Experiment 3 utilized two measures of parsing that in past research have provided evidence to distinguish processing preferences across German and English adults: word length and body neighborhood size, the number of words that have the same orthographic rime as the target stimulus. As mentioned in Chapter 1, a study by Ziegler et al. (2001) examined these two variables with native speakers of English and German. In the Ziegler et al. study, both groups named words and nonwords in their L1 that differed along these two dimensions. The results of this study demonstrated interactions between language and word length, and language and number of body neighbors. The interaction between language and length occurred because German participants naming words in German were affected more by word length than English speaking participants naming words in English. On average, German participants were slower to name 6 letter words than 3 letter words. In contrast, the language by number of body neighbors interaction was due to larger effects of body neighborhood size for English participants. On average, words from large body neighborhoods were named more quickly by English speakers than words from small body neighborhoods. For the native German speakers, the difference for body neighborhood size was in the same direction, but only marginal.

What is remarkable about the findings from the Ziegler et al. (2001) study is that not only there were differences in performance for the two language groups, but these differences were found despite the striking similarities in the stimulus items in both languages. The critical items in the Ziegler et al. experiment were English-German cognates, translations that are identical or nearly identical in form and meaning in German and English (e.g., hand-*Hand*). The observed differences between the two native speaker groups can therefore not be easily attributed to differences in stimulus materials, but rather to the processing strategies adopted by participants with different native languages.

The current study utilized the same methods and items as Ziegler et al. (2001). However, German-English bilinguals and native English speakers named the English version of the cognate materials (see Appendix C). The goal was to determine whether the pattern of results for non-native speakers of English would be similar to those for native speakers, as the previous two experiments suggest, or whether the German speakers would produce the distinct effects reported by Ziegler et al. even when they are naming words in their L2. In addition, the use of cognates in the current study should increase the probability that German will be active while naming. If native German participants do not differ from native English participants for this task, this result would provide strong evidence that this highly proficient native speaker group was able to apply L2 processing preferences even when the L1 is likely to be highly active (e.g. Dijkstra & Van Heuven, 2002).

Method

Participants

As in the Ziegler et al. (2001) study, the two participant groups of interest were native speakers of English (n=23) and late German-English bilinguals (n=54). The native German speakers in this study were the same participants as in Experiment 2. They were recruited from the University of Muenster and the Max Planck Institute of Cognitive Neuroscience in Leipzig and were paid for their participation. The native speakers of English were recruited from The Pennsylvania State University through the psychology department subject pool and received course credit for their participation.

Materials

The materials were the same English word and nonword materials used by Ziegler et al. (2001). The words and nonwords varied both by length in that they were either 3, 4, 5, or 6 letters long, and also by number of body neighbors (body N). All words were either English-German cognates that were identical or similar in form across both languages or words that were not cognates but similar in phonological onset to a word of the same length in German[ii].

Procedure

Participants completed the naming task one at a time at a computer. On each trial, a 500 ms fixation appeared on the screen followed by a string of letters. Participants responded by saying the string of letters aloud into a voice key which recorded the onset of naming. After the voice key was activated the word disappeared. It was then followed by the appearance of a fixation signaling the start of the next trial. Participants were instructed to name the string as quickly and as accurately as possible and to pronounce the nonwords as if they were real English words. In addition, participants were asked to be cautious about making any extraneous noises that might activate the voice key prematurely. The 160 words and nonwords from this study were randomly divided among four blocks to allow participants to take a short break if necessary. Within each block, words and nonwords were presented in random order. Naming responses were tape recorded and these recordings were later used to code participant responses for accuracy.

Results and Discussion

Because adult cross linguistic research has demonstrated significant differences in the L1 performance for both of these participant groups in English, naming latencies and accuracy for native English and native German speakers were compared to determine whether manipulating word length and number of body neighbors revealed differences in processing preferences across these two groups. Separate analyses were conducted for word and nonword items.

Exclusion of participants

Participants who did not meet certain language and accuracy criteria were excluded from the analyses. As in previous experiments, participants were excluded if they were not native speakers of either English or German. For the English group, five participants were excluded from the analysis and for the German native speaker group three participants were excluded from the analysis. Participants were also excluded if less than 80% of the word and nonword data were not correct, non voice-key error trials[iii]. Only one participant in the native English group failed to meet these criteria, however 18 participants from the native German group were excluded due to low accuracy. In addition, one other participant was excluded from the native English group. This participant had over five years of German and was both a second language major and minor at Penn State. During the course of the experiment, pronunciations for nonwords seemed somewhat Germanized, and this participant even gave Germanized pronunciations to a small number of English words (i.e. WINE *vine*). After these exclusions, 33 native German speakers, and 17 native English speakers were included in the final analyses.

Calculation of outliers

For each participant, outlier trials were identified by finding their grand mean and the reaction time 2.5 standard deviations above and below that mean. Words and nonwords falling outside of that range were not included in the following analyses. Outlier trials counted for less than 3% of the data.

Native English and Native German body x length comparison: word RT

In this analysis, reaction time to words was examined using a 2 (body N) X 4 (word length) X 2 (language group) ANOVA. In the participants analysis, body N and

Table 5.1

Reaction Time (ms) for words by body neighborhood size and word length for native English speakers, native German Speakers, and Ziegler et al. (2001)

Conditions		Participant Group			
Body N Size	Word Length	Native English	Native German	Ziegler et al. (2001) English Data	Ziegler et al. (2001) German Data
Small body N	3	606 (23.6)	637 (17.0)	526	524
	4	633 (26.3)	672 (18.9)	527	521
	5	662 (22.6)	644 (16.2)	528	566
	6	656 (25.6)	702 (18.4)	534	563
Large body N	3	588 (22.4)	659 (16.1)	508	512
	4	640 (26.9)	669 (19.3)	505	501
	5	638 (23.5)	624 (16.9)	515	568
	6	640 (27.8)	664 (19.9)	526	572

Standard Errors in parentheses

word length were within-participants variables and language group was a between-participants variable. For the item analysis, language group was a within-item variable, and body N and word length were between-items variables.

This analysis revealed a main effect of body N that was significant by participants ($F1(1,48)=11.94$, $MSE = 225.5$, $p < 0.002$), but only marginal by items ($F2(1,72)=3.36$, $MSE = 2166.2$, $p < 0.08$). In addition, there was a main effect for word length that was significant both by participants ($F1(3,144)=16.49$, $MSE = 1818.6$, $p < 0.001$), and by items ($F2(3,72)=5.86$, $MSE = 2166.2$, $p < 0.002$). The main effect for body N was due to significantly faster response times for large body N words compared to small body N words by an average of 11 ms. For word length, after Bonferroni correction, pairwise comparisons revealed that response times to 3 letter words were significantly faster than 4 letter words ($t_{49} = 5.72$, $p < .001$) and 6 letter words ($t_{49} = 6.54$, $p < .001$). Surprisingly

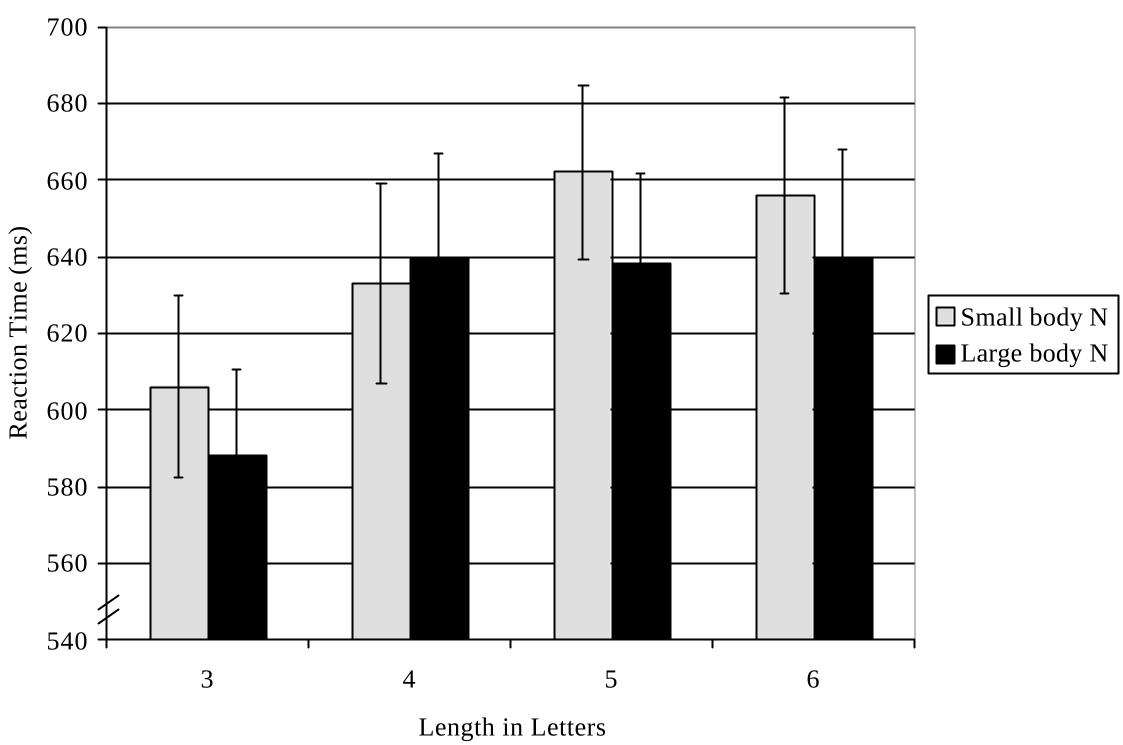


Figure 5.1. Reaction time (ms) for words for body N size by length for the native English speaker group.

response times for 4 letter words were significantly slower than for 5 letter words (t_{49} = 2.94, $p < .03$), and finally 5 letter words were significantly faster than 6 letter words (t_{49} = 4.14, $p < .001$).

These two main effects were qualified by a significant interaction between body N and length that was significant by participants (*F1*(3,144)=5.57, *MSE* = 1000.9, p < 0.002), but not by items (*F2*(3,72)=1.23, *MSE* = 2166.2, p < 0.4). Post hoc analyses of this interaction suggest that the effects of word length was stronger in the small body N condition (*F1*(3,147)=21.83, *MSE* = 1426.7, p < 0.007) than the large body N condition (*F1*(3,147)=6.21, *MSE* = 1815.0, p <0.002).

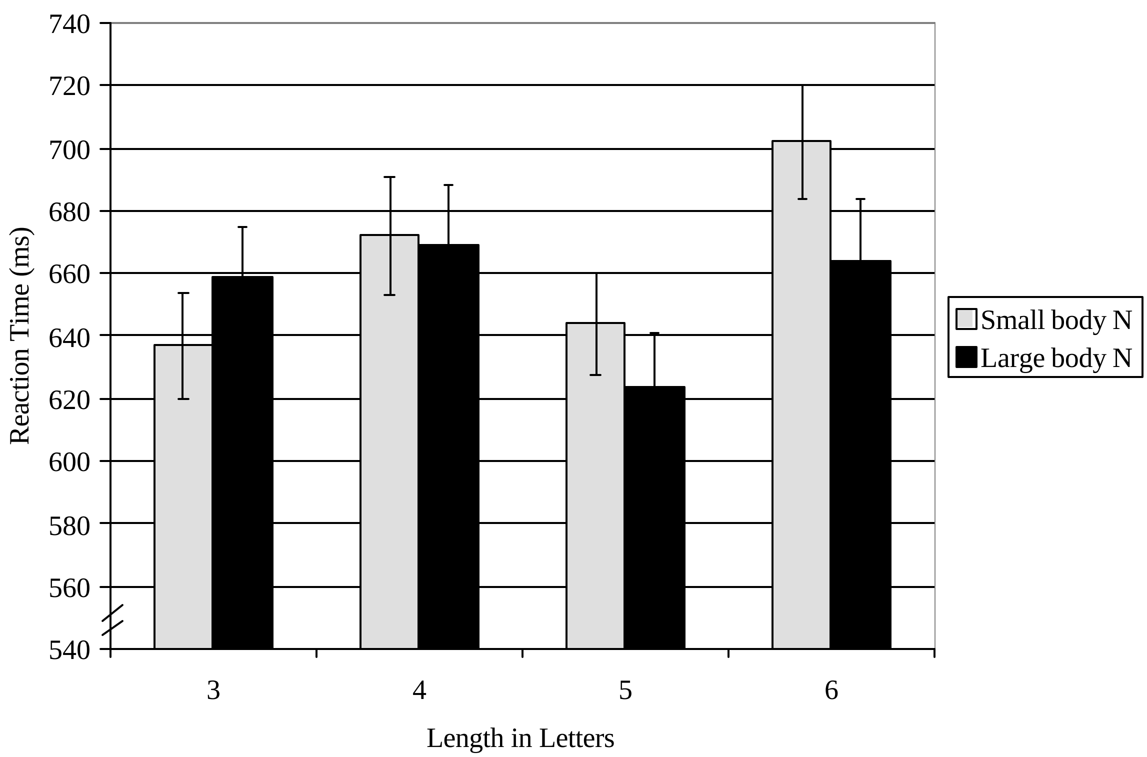


Figure 5.2. Reaction time (ms) for words for body N size by length for the native German speaker group.

The analysis for words also revealed an overall effect for language group that was not significant by participants (*F1*(1,48)=0.80, *MSE* = 75118.5, $p < 0.4$), but was significant by items (*F2*(1,72)=16.36, *MSE* = 1348.2, $p < 0.001$). This main effect by participants was due to overall faster response times for words for native English speakers over native German speakers by an average of 24 ms.

In addition, interactions with language qualified a number of the findings reported earlier. There was an interaction of language and word length that was significant by participants (*F1*(3,144)=4.23, *MSE* = 1818.6, $p < 0.001$), and by items (*F2*(3,72)=6.27, *MSE* = 1348.2, $p < 0.001$). This interaction occurred because for native English speakers, 3 letter words were significantly faster than 4, 5, and 6 letter words, and for native German speakers, the pattern followed the results of the overall effect for word length. Three letter words were significantly faster than 4 and 6 letter words. Four letter words were significantly slower than 5 letter words, and 5 letter words were significantly faster than 6 letter words.

Most critically, these effects were qualified by a three way interaction between language group, word length, and body N. This interaction was significant by participants, (*F1*(3,144)=4.23, *MSE* = 1000.9, $p < 0.007$), but not by items (*F2*(3,72)=0.38, *MSE* = 1348.2, $p < 0.4$). This interaction revealed that the body N by word length interaction was significant only for the native German participants (*F*(3,96)=9.59, *MSE* = 1181.9, $p < 0.001$), and not the native English participants (*F2*(3,48)=2.49, *MSE* = 638.8, $p < 0.2$).

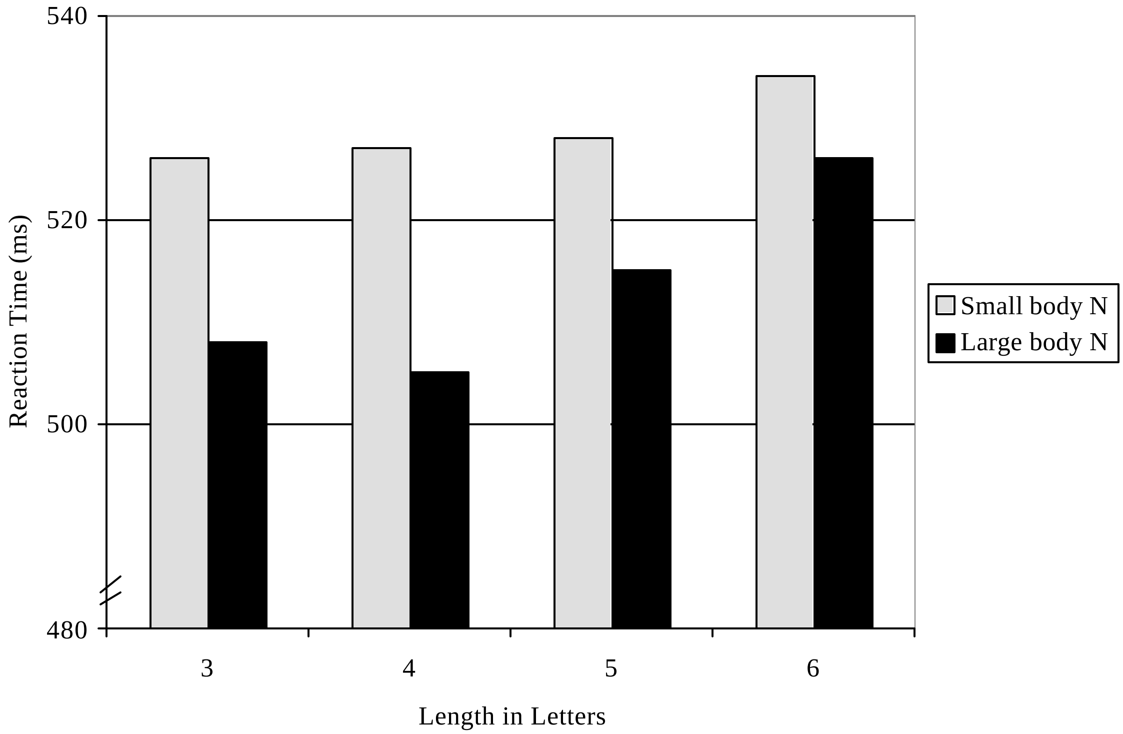


Figure 5.3. Reaction time (ms) for words for body N size by length for English participants (Ziegler et al., 2001)

The results of this analysis suggest that when naming real English words, the native German participants are relying more on L2 processing strategies than strategies from their L1. This seems to be the case because the German participants in the Ziegler et al. (2001) study that named the German version of this stimulus list did not show a body length effect for German, only a length effect. In this experiment, both native German speakers and native English speakers responded to large body N words faster than small body N words suggesting that both groups were processing the orthographic rime. However the results also suggest that there was not complete overlap in processing

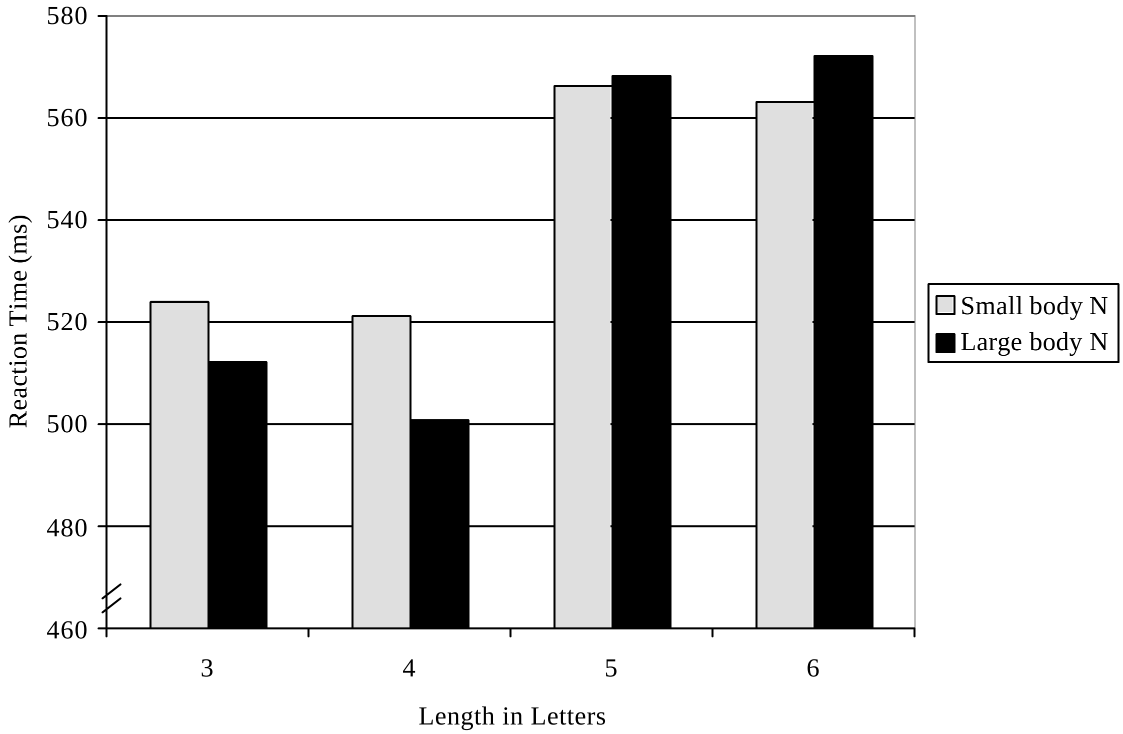


Figure 5.4. Reaction time (ms) for words for body N size by length for German participants (Ziegler et al., 2001)

between these two groups. The three-way interaction between language, length, and body N size demonstrated that length effects for native German speakers were stronger for small body N words than for large body N words. For native English speakers, word length and body N did not interact. For the native German group, this interaction may have occurred because the large body N words may have seemed more representative of English words than the small body N words, thus making larger unit processing more likely. In addition, the words in the small body N condition may have seemed more similar to German words than the words in the large body N condition. In this condition 14 of the 40 words were cognates that were identical in form between English and

German (e.g., FORM-*FORM*). For the large body N condition, only 8 of the 40 words were identical form cognates. In order to determine if the effects mentioned earlier exist despite these possible cross-language lexicality differences, analyses were conducted examining the effects of length and body N for nonwords.

Native English and Native German body x length comparison: nonword RT

For nonwords, the three way interaction between language, word length, and body N size was not significant by participants, ($F1(3,144)=1.02$, $MSE = 1922.8$, $p < 0.4$), or by items ($F2(3,72)=0.44$, $MSE = 1459.3$, $p < 0.8$). The overall effect of language again was not significant by participants ($F1(1,48)=0.27$, $MSE = 121179.6$, $p < 0.7$), but was significant by items ($F2(1,72)=6.26$, $MSE = 1459.3$, $p < 0.02$). For nonwords the difference between native English and native German participants was on average 15 ms. Although there was an overall effect of language by participants, language only marginally interacted with body N by participants ($F1(1,48)=3.46$, $MSE = 1298.6$, $p < 0.07$), and by items ($F2(3,72)=2.87$, $MSE = 1459.3$, $p < 0.1$). This marginal interaction was due to a larger naming advantage for large body N nonwords for native English speakers than native German speakers. The average advantage for native English speakers was 25 ms and for the native German group, the average advantage was only 11 ms.

In addition to these somewhat tenuous effects by language, the nonword analysis revealed significant main effects for word length and for body N size. The effect of nonword length was significant by participants ($F1(3,144)=19.92$, $MSE = 3025.0$, $p < 0.001$), and by items ($F2(3,72)=8.24$, $MSE = 3756.0$, $p < 0.001$). Pairwise comparisons revealed a number of significant differences between the four length conditions. Three letter nonwords were named significantly faster than 4 letter nonwords (t_{49} = 4.33, p < .001), 5 letter nonwords (t_{49} = 5.28, p < .001), and 6 letter nonwords (t_{49} = 5.38, p < .001). Four letter nonwords were named faster than 5 letter nonwords (t_{49} = 2.93, p < .03), and 6 letter nonwords (t_{49} = 3.50, p < .006).

Table 5.2

Reaction Time (ms) for nonwords by body neighborhood size and word length for native English speakers, native German Speakers, and Ziegler et al. (2001)

Conditions		Participant Group			
Body N size	Word Length	Native English	Native German	Ziegler et al. (2001) English Data	Ziegler et al. (2001) German Data
Small body N	3	653 (27.4)	679 (19.7)	578	568
	4	669 (29.9)	698 (21.5)	618	588
	5	729 (36.6)	720 (26.3)	623	637
	6	737 (35.2)	737 (25.3)	654	657
Large body N	3	632 (29.2)	676 (21.0)	568	563
	4	681 (30.4)	701 (21.8)	588	565
	5	681 (29.0)	708 (20.8)	622	621
	6	692 (35.2)	707 (25.3)	628	687

Standard Errors in parentheses

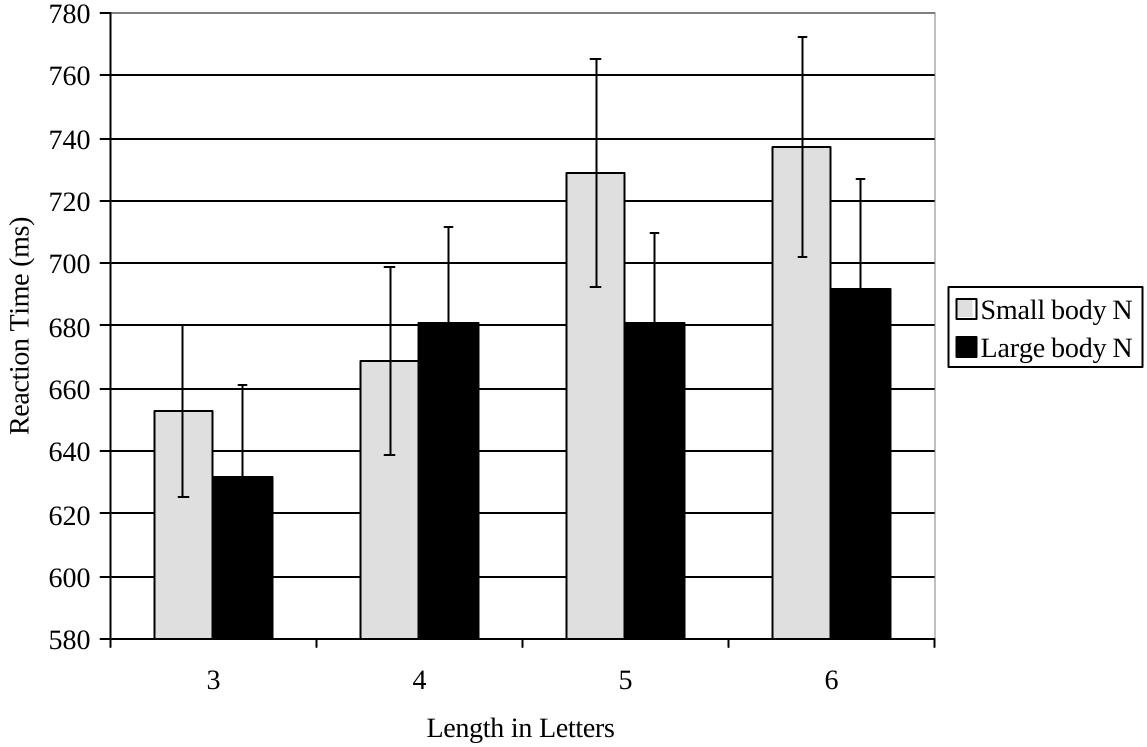


Figure 5.5. Reaction time (ms) for nonwords for body N size by length for the native English speaker group.

For nonwords the effect of body N size was also significant by participants (*F1*(1,48)=22.28, *MSE* = 1298.6, $p < 0.001$), and marginal by items (*F2*(3,72)=8.24, *MSE* = 3756.0, $p < 0.001$). As in the word analysis, this main effect for word body was due to faster naming latencies for large body N nonwords than small body N nonwords by an average of 18 ms.

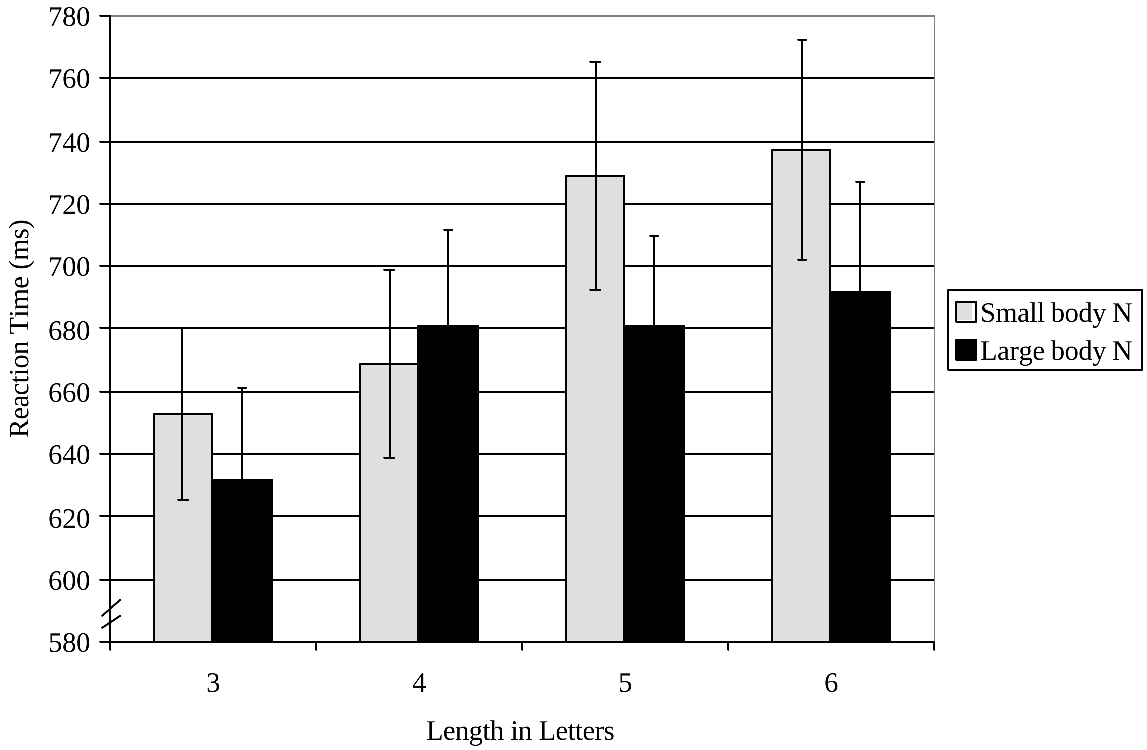


Figure 5.6. Reaction time (ms) for nonwords for body N size by length for the native German speaker group.

The findings from the nonword analysis suggest that naming strategies for nonwords across both native English speakers and German-English bilinguals were similar. There was some evidence that the advantage for large body N words was greater for the native English group than for the native German group, consistent with the results of the cross-linguistic comparison in the Ziegler et al. (2001) study, but the group differences were not significant.

The results also provide some evidence that the interactions with language revealed in the word analysis may have been due to the presence of more identical form

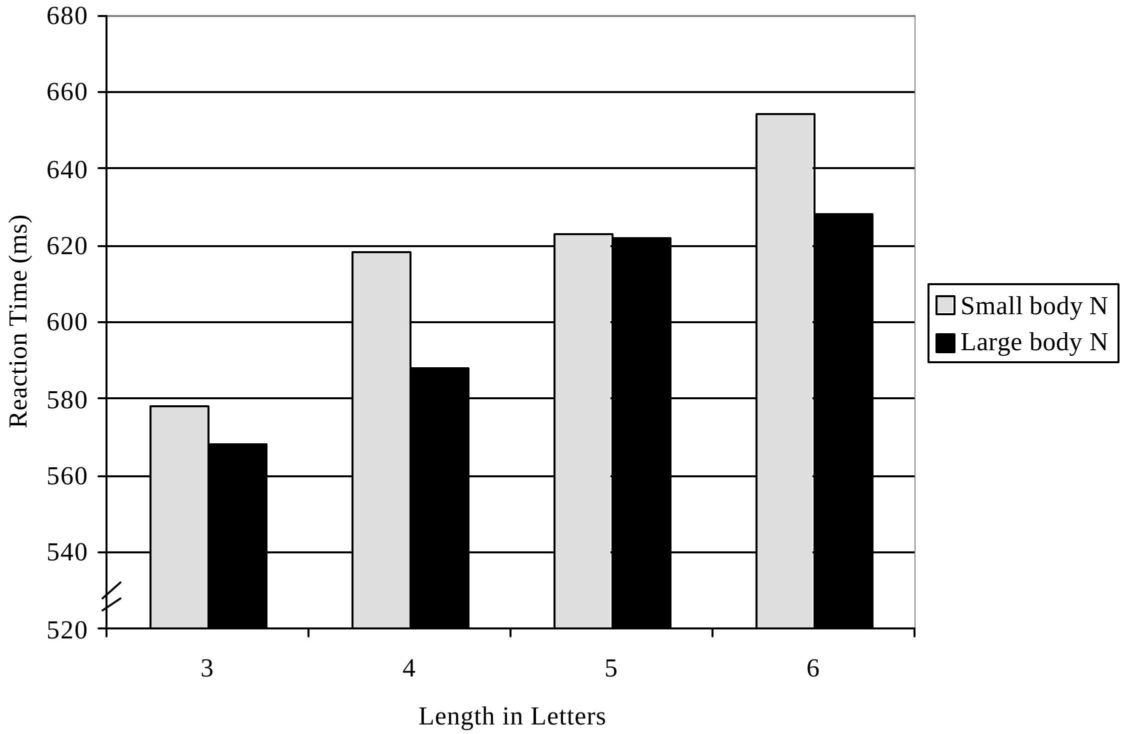


Figure 5.7. Reaction time (ms) for nonwords for body N size by length for English participants (Ziegler et al., 2001).

cognates for the low body N group than for the high body N group. The presence of a greater number of identical form cognates combined with lower body N for English, may have allowed German L1 processing strategies to emerge for these stimuli. This point will be revisited at the end of this chapter.

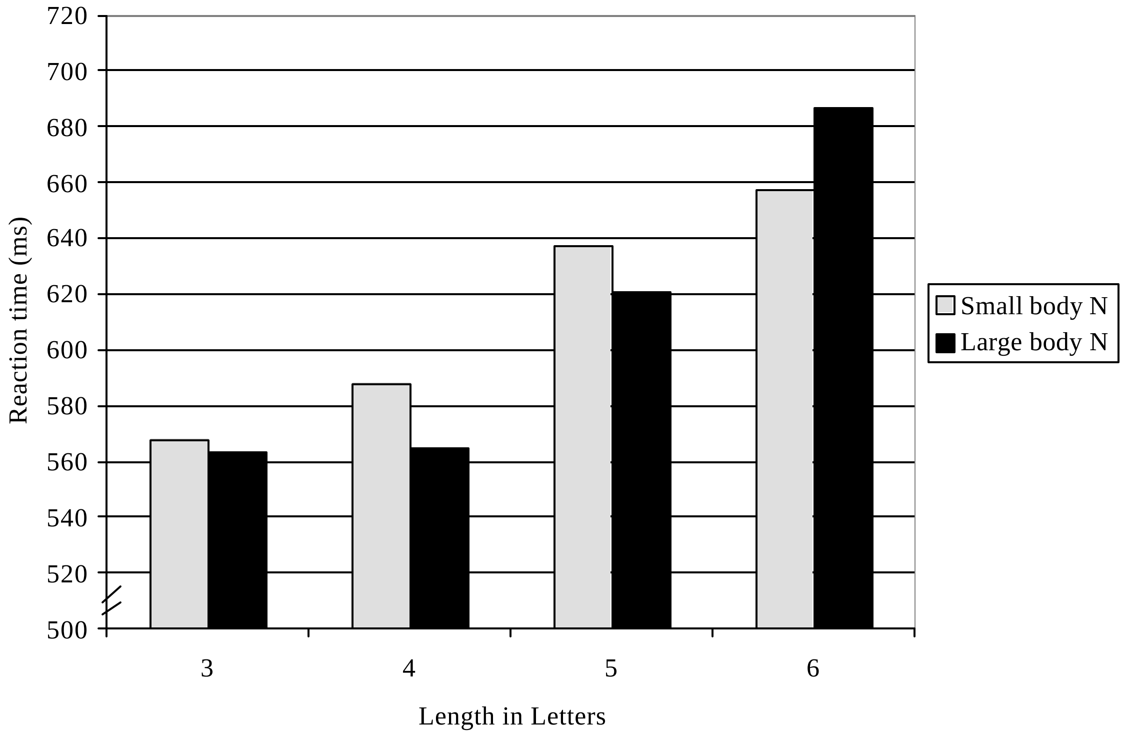


Figure 5.8. Reaction time (ms) for nonwords for body N size by length for German participants (Ziegler et al., 2001).

Native English and Native German body x length comparison: word accuracy

Although Ziegler et al. (2001) did not analyze the accuracy of their participants in terms of the experimental variables, it is possible that the error data may shed light on either parsing strategies that were utilized by each participant group or on general task strategies that may be influencing performance on the experimental variables. For the following analyses, word and nonword data were analyzed separately. As in the reaction time analyses, for the participant ANOVA, length and body N size were within-participant variables and language group was a between-participants variable. For the item analyses, language group became a within-item variable and length and body N size were between-item variables.

A 2 (body N size) X 4 (length) X 2 (language group) ANOVA was used to analyze the accuracy for words. This analysis revealed a main effect of length that was significant by participants ($F1(3,144)=4.58$, $MSE = .009$, $p < 0.005$), but not significant by items ($F2(3,72)=2.12$, $MSE = .009$, $p < 0.2$). There was a main effect for language that was significant by participants ($F1(1,48)=7.12$, $MSE = .02$, $p < 0.02$), and by items ($F2(1,72)=10.21$, $MSE = .005$, $p < 0.003$). On average, native English speakers had higher naming accuracy than native German speakers by an average of 3.6%.

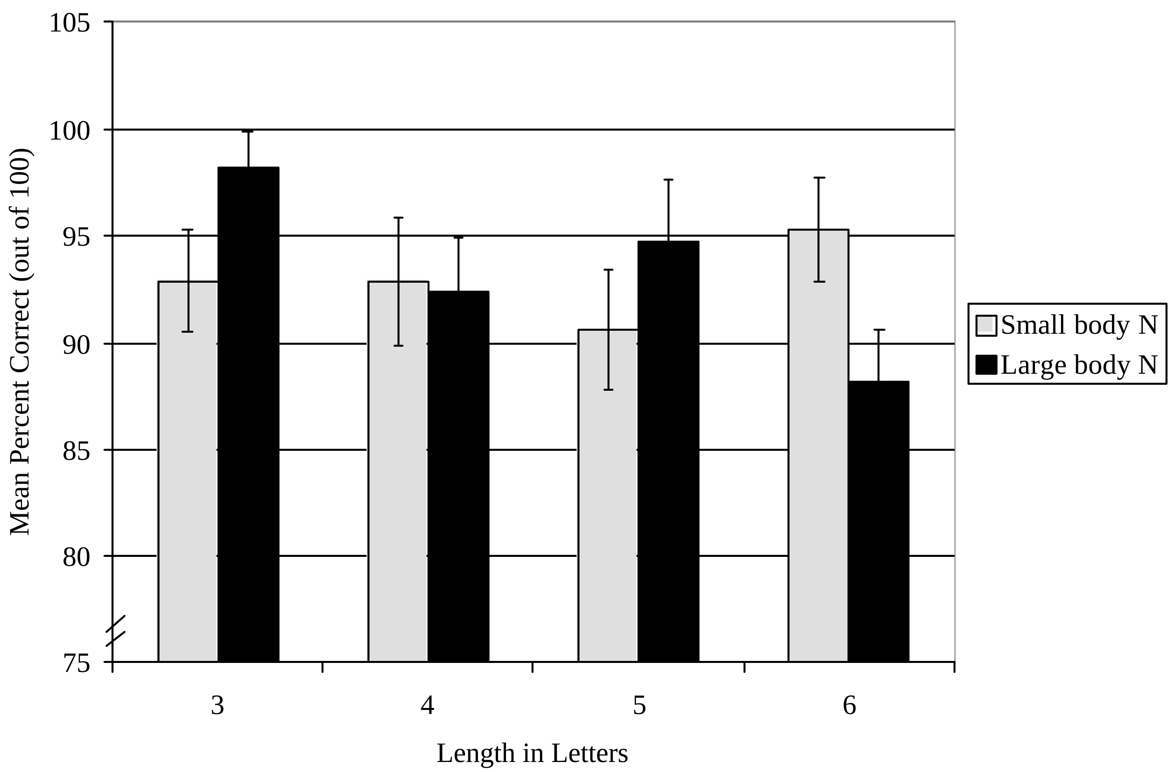


Figure 5.9. Accuracy for words for body N size by length for Native English speakers.

Although there was a main effect for language, no interactions between language, body N size, and word length reached significance. The only significant interaction was the interaction between body N size and length. This interaction was significant by participants (*F1*(3,144)=5.04, *MSE* = .009, $p < 0.003$), and marginal by items

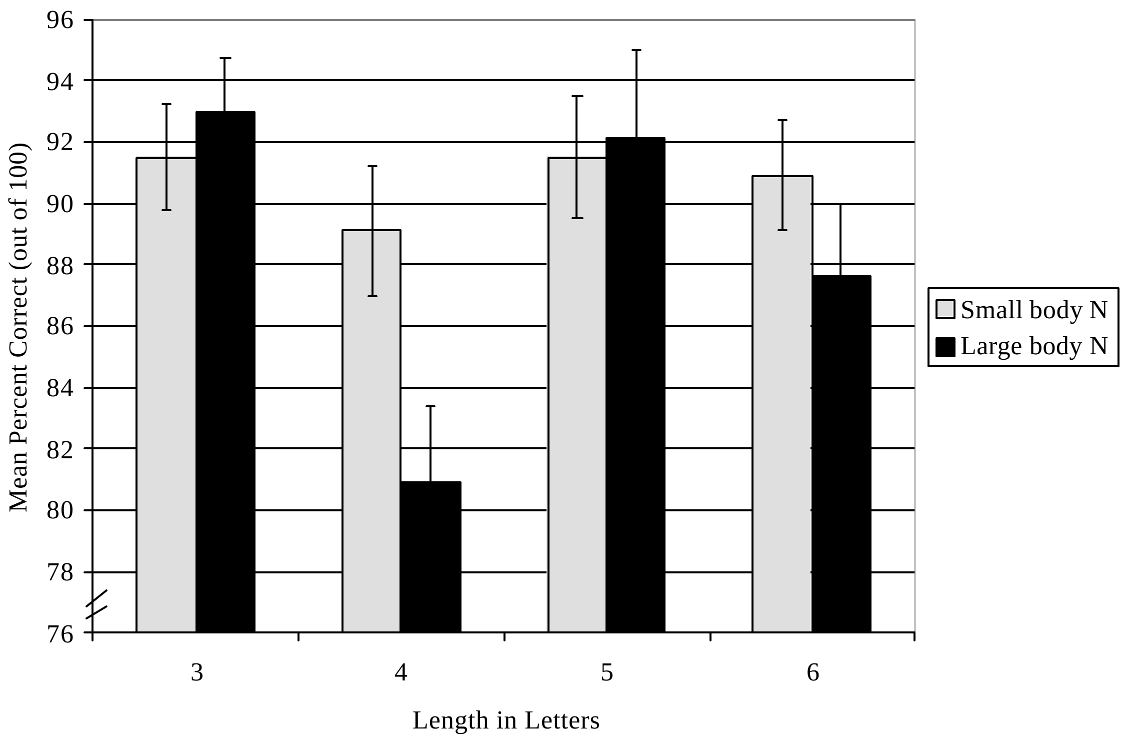


Figure 5.10. Accuracy for words for body N size by length for native German speakers.

(*F2*(3,72)=2.18, *MSE* = .009, $p < 0.1$). This interaction was due to significant effects of length for large body N words (*F*(3,147)=12.32, *MSE* = .009, $p < 0.001$), but not small body N words (*F*(3,147)=0.39, *MSE* = .01, $p < 0.8$). For large body N words, there were significant accuracy differences for 3 vs. 4 letter words, 3 vs. 6 letter words, and 4 vs. 5 letter words. On average, 3 letter words were 10% more accurate than 4 letter words ($t_{49} = 6.19$, $p < .001$), and 7% more accurate than 6 letter words ($t_{49} = 3.73$, $p < .006$). In addition, 5 letter words were 8.2 more accurate than 4 letter words ($t_{49} = 4.07$, $p < .001$), and 5.2 % more accurate than 6 letter words ($t_{49} = 2.74$, $p < .009$).

Table 5.3

Accuracy (out of 100%) for words by body neighborhood size and word length for native English speakers and native German Speakers

Conditions		Participant Group	
Body N Size	Word Length	Native English	Native German
Small body N	3	92.9	91.5
		(2.4)	(1.7)
	4	92.9	89.1
		(3.0)	(2.1)
	5	90.6	91.5
		(2.8)	(2.0)
	6	95.3	90.9
		(2.4)	(1.8)
Large body N	3	98.2	93.0
		(1.7)	(1.2)
	4	92.4	80.9
		(2.5)	(1.8)
	5	94.7	92.1
		(29.0)	(1.5)
	6	88.2	87.6
		(2.4)	(1.8)

Standard Errors in parentheses

The analysis of word accuracy failed to find differences between the native English group and the native German group. Although overall native English participants were more accurate than participants in the native German group, language group failed to interact with any of the main experimental variables, demonstrating that the main effect of length and the interaction between body N and length were similar across both groups.

Native English and Native German body x length comparison: nonword accuracy

Although there were no differences between the two groups in the analysis by word accuracy, the analysis for nonword accuracy did reveal significant cross-language group differences. This analysis revealed main effects for body N size and length. The main effect for body N size was significant by participants (*F1*(1,48)=22.91, *MSE* = .01, $p < 0.001$), and by items (*F2*(1,72)=7.00, *MSE* = .02, $p < 0.02$). The main effect was due to greater accuracy for onset-rime parsed nonwords over nononset parsed nonwords by an average of 5.5%.

The main effect for length was significant by participants (*F1*(3,144)=5.27, *MSE* = .01, $p < 0.002$), but not by items (*F2*(3,72)=1.43, *MSE* = .02, $p < 0.3$). However after Bonferroni correction, no pairwise comparisons reached significance. These two main effects however were qualified by a significant interaction between body N size and length that was significant by participants (*F1*(3,144)=4.78, *MSE* = .01, $p < 0.004$), but not by items (*F2*(3,72)=1.37, *MSE* = .02, $p < 0.3$). The interaction by participants was due to significant differences between 5 and 6 letter nonwords only for small body N nonwords ($t_{49} = 3.42$, $p < .02$) by an average of 8%.

There was no overall effect of language group by participants (*F1*(1,48)=0.81, *MSE* = .02, $p < 0.4$), or by items (*F2*(1,72)=0.71, *MSE* = .009, $p < 0.5$). Language did interact with nonword length and also with both nonword length and body N size. The interaction of language group and nonword length was significant by participants

(*F1*(3,144)=4.43, *MSE* = .01, p < 0.006), but only marginal by items (*F2*(3,72)=2.32, *MSE* = .009, p < 0.09). This interaction was significant by participants because there were no accuracy differences between nonword lengths for native German speakers. However for native English speakers, 5 letter phonemes were less accurate than both 4 letter nonwords ($t_{49} = 3.56$, $p < .04$) and 6 letter nonwords ($t_{49} = 4.40$, $p < .005$).

Table 5.4

Accuracy (out of 100%) for nonwords by body neighborhood size and word length for native English speakers and native German Speakers

Conditions		Participant Group	
Body N size	Word Length	Native English	Native German
Small body N	3	87.6	86.1
		(2.9)	(2.1)
	4	89.4	81.2
		(2.8)	(2.0)
	5	71.2	83.0
		(3.1)	(2.3)
	6	89.4	85.8
		(2.6)	(1.9)
Large body N	3	90.6	88.5
		(2.4)	(1.7)
	4	91.2	90.3
		(2.3)	(1.7)
	5	90.0	88.8
		(2.1)	(1.5)
	6	91.2	87.0
		(2.4)	(1.7)

Standard Errors in parentheses

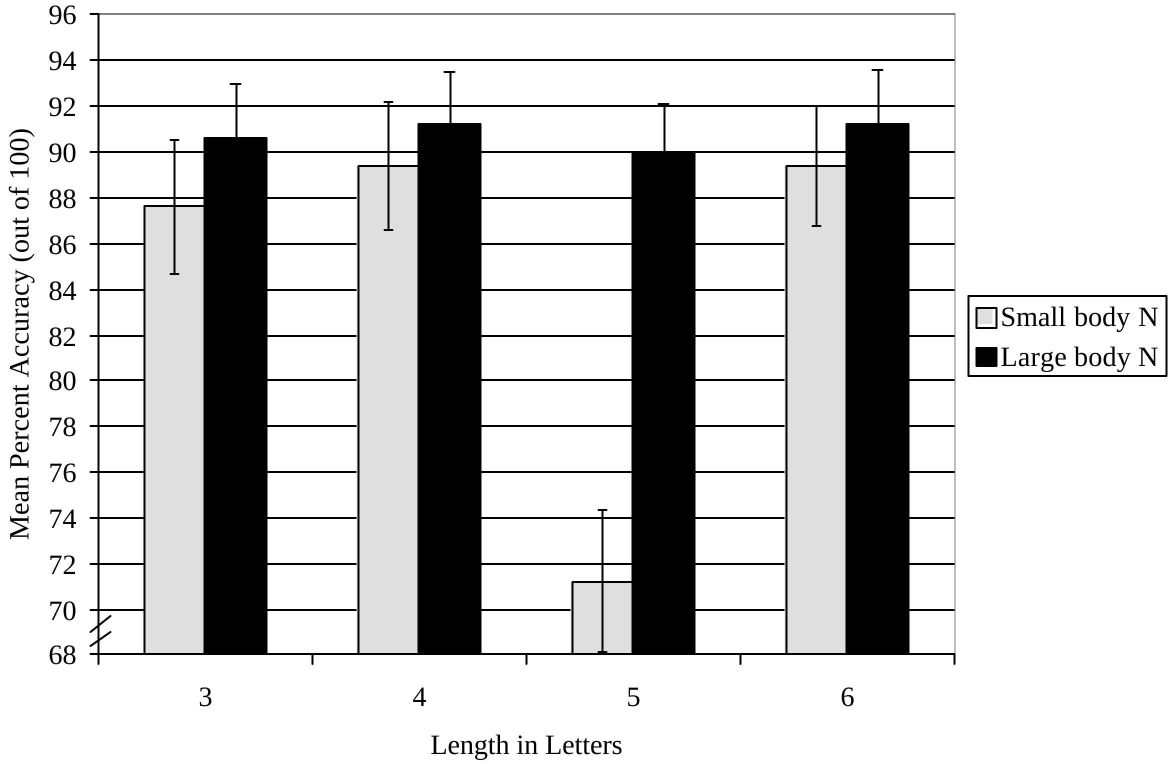


Figure 5.11. Accuracy for nonwords for body N size by length for native English speakers.

Finally the three-way interaction between language group, body N size, and nonword length was significant by participants (*F1*(3,144)=3.64, *MSE* = .01, $p < 0.02$), but not by items (*F2*(3,72)=2.02, *MSE* = .009, $p < 0.2$). This interaction was significant by participants because the interaction between body and length was significant for the native English participants (*F*(3,48)=4.15, *MSE* = .01, $p < 0.02$), but only marginal for the native German participants (*F*(3,96)=2.23, *MSE* = .009, $p < 0.1$).

Although the nonword accuracy data differ for these two groups, it is difficult to determine whether these differences indicate different processing strategies or simply differences in language knowledge for the two groups. One similarity for both groups is that large body nonwords had high accuracy. So both the native English and the native German participants received a benefit for naming words that had familiar English rimes.

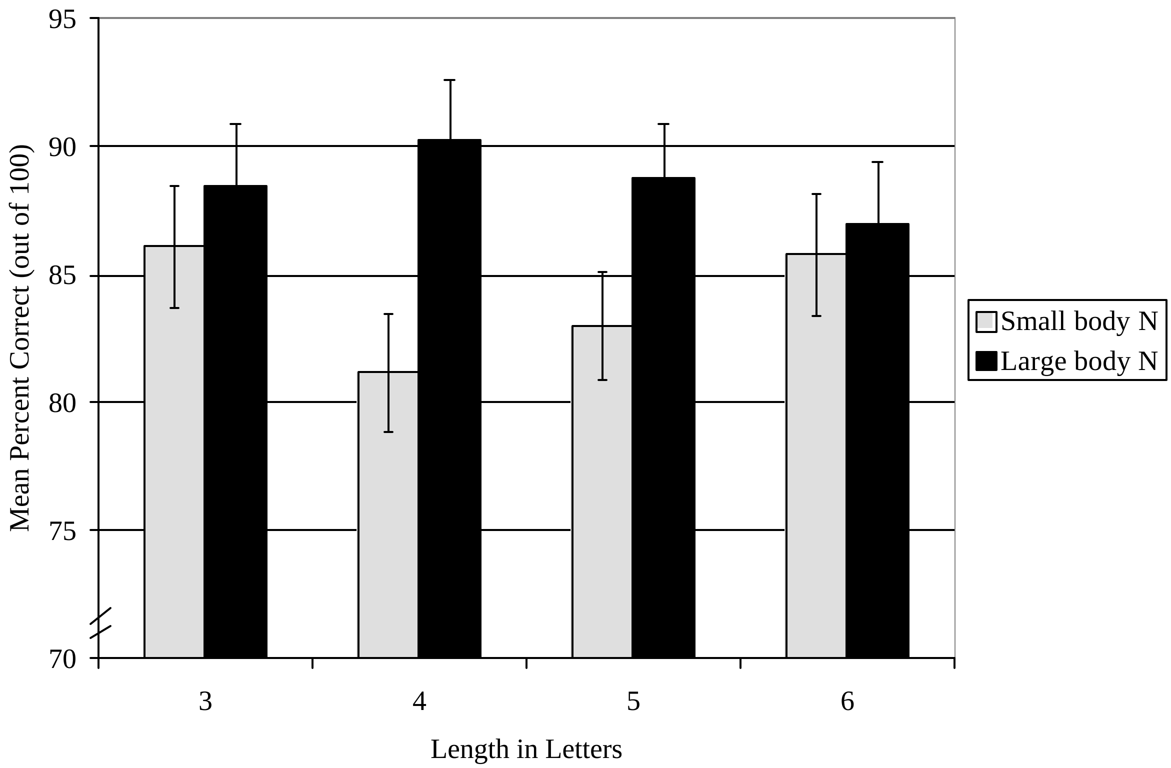


Figure 5.12. Accuracy for nonwords for body N size by length for native German speakers.

General Discussion

As in Experiments 1 and 2, the general pattern of findings for native German speakers for the English naming task is more similar to the native English speakers performing naming in English, than to the native German speakers naming similar words in German from the Ziegler et al. (2001) study. If native German speakers were using L1 processing strategies while naming English words and nonwords, they would have exhibited stronger length effects than native English speakers for words and especially for nonwords where naming generally relies more on computed verses directly retrieved phonology. However, for nonwords, the length effect for native English speakers and native German speakers is more similar than that of the native German participants in the Ziegler et al. (2001) study.

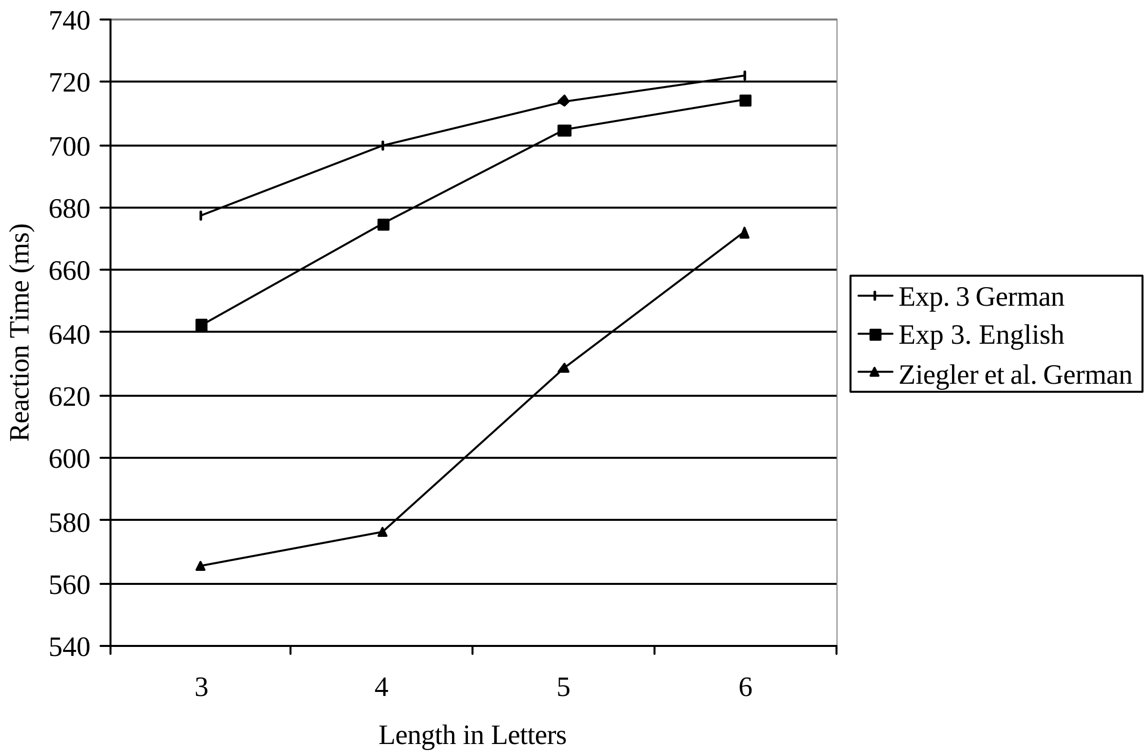
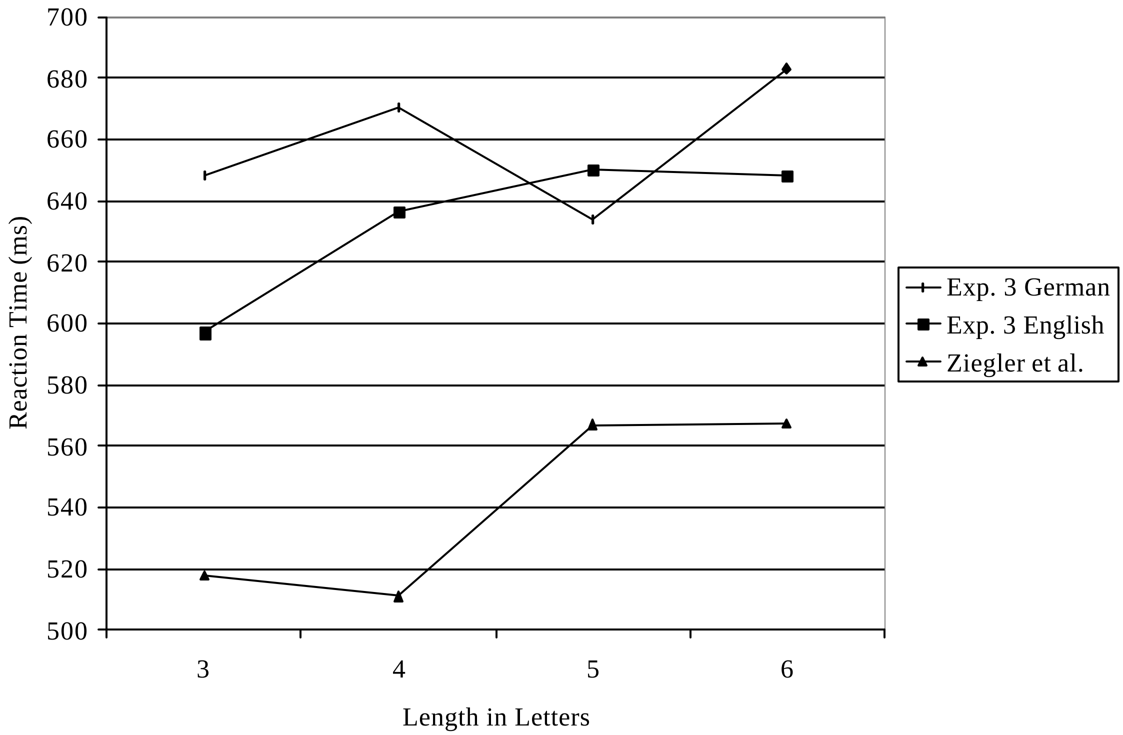


Figure 5.13. Length effect for nonwords in English in Experiment 3 and for German in Ziegler et al. (2001).

However unlike the previous lexical decision experiments, the naming task revealed some distinctive findings between these two groups. The length effect for native English and native German speakers did differ. Yet the effect in the native German speakers appeared to be distinct from the German speakers in Ziegler et al. (2001) as well. For naming words, German native speakers showed larger length effects for small body N words than for large body N words. In contrast, native English speakers did not show this pattern. The stronger length effects for the small body N words may have been



<u>Figure 5.14.</u> Length effect for words in English in Experiment 3 and for German in Ziegler et al. (2001).

due to the greater similarities to words in this condition to German words than the words in the large body condition. As mentioned earlier, the small body N condition had more identical German-English cognates than the large body N condition. The larger L1 cue from words in this condition, may have influenced the processing and therefore allowed a greater reliance on L1 processing strategies than the lexical decision tasks in Experiments 1 and 2. As shown in Figure 5.15, this appears to be the case. The length effect in the low body N words is more similar to the native German group than the length effect in the high body N words. This interaction did not occur for nonwords, however, suggesting that the higher similarity to German of the low body N cognates may have been driving the interaction for words. For nonwords, both native English speakers and German-English bilinguals showed benefits for naming words with large numbers of body neighbors compared with smaller numbers of body neighbors.

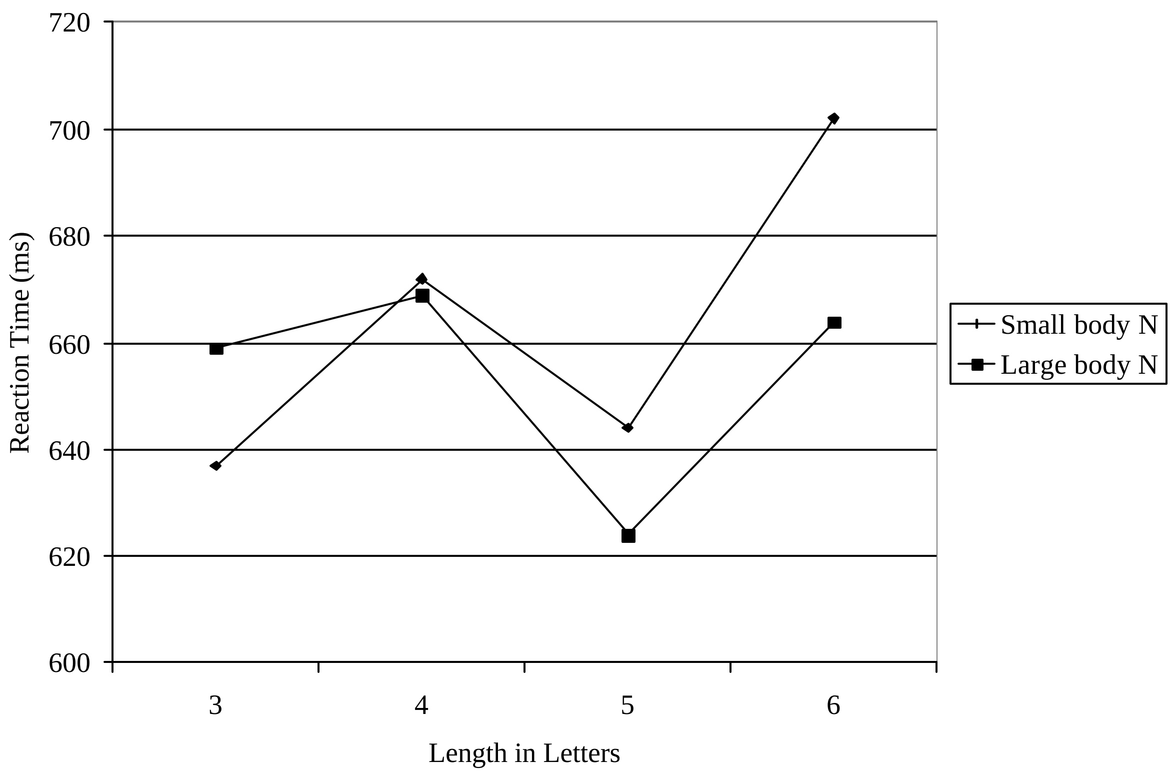


Figure 5.15. Length by body N interaction for native German speakers in Experiment 3.

Because nonwords have often been used to measure how spelling-to-sound correspondences are computed, the fact that native German speakers showed similar response patterns to English native speakers for nonwords suggests that these proficient bilinguals may be using the same strategies to map spelling-to-sound as the native English speakers. However, the presence of the interaction of body N size and length for the word items does suggest that although the German-English bilinguals are highly proficient in English and in general able to use native like processing strategies, the first language is active and ready to compete for selection.

Overall, in this Experiment, proficient German-English bilinguals were more similar to the native English participants than to the German participants in the Ziegler et al. (2001) study, suggesting that they had access to L2 processing strategies. When English nonwords had many body neighbors, the processing of these stimuli was very similar to native speakers of English. However for the small body N words, items that collectively had more identical cognates, word length factored into how quickly these items were processes suggesting that these speakers were relying somewhat on L1 lexical processing strategies. Although it would have been ideal to have a group of German-English bilinguals completing the German version of this task to provide a more complete comparison, the fact that the participants in the Ziegler et al. study were most likely proficient bilinguals suggests that the processing revealed in this study was due to L2 processing strategies and not a general processing strategy adapted for both languages. In general this pattern of results in addition to the general length effects for both groups is consistent with a Dual-Route model where nonnative speakers are able to acquire some of the processing rules of the L2. However, even for highly proficient bilinguals, these rules are weaker and in the presence of strong L1 cues, L1 strategies emerge. This point will be revisited in the final chapter.

Chapter 6: General Discussion

Cross-linguistic research has demonstrated differences in the lexical processing strategies adopted by native speakers whose languages vary in orthographic depth. The three experiments reported here used methods drawn from cross-linguistic research on lexical parsing to determine whether bilinguals who speak English as a second language use the same parsing strategies as native English speakers or transfer aspects of their first language preferences to the second language. In Experiment 1, the performance of German-English and Japanese-English bilinguals was compared to that of native English speakers. In Experiments 2 and 3, the German-English bilinguals were compared to native English speakers. In each experiment, the target language was always English. What varied was whether English was the L1 or L2 for participants. In Experiments 1 and 2, using lexical decision tasks, few differences were found for proficient bilinguals in their L2 and native English speakers in their L1, suggesting that it is possible for non-native speakers to acquire the distinct processing strategies associated with the L2. When the task was changed to word naming in Experiment 3, and the salience of the L1 was increased by using cognate materials, group differences emerged, suggesting that even highly proficient bilinguals who are able to adopt the L2 strategies under some circumstances also fall back into their preferred L1 strategy when the cues are consistent with the L1. In the next section, we first summarize the major findings of each of the experiments and then consider the implications of the results for the lexical models described in the introduction.

Summary of findings

In Experiment 1, three groups of participants completed a lexical decision task in which nonwords were either pseudohomophones, neighbors of the pseudohomophone's base word (O+P+ controls) or nonwords unrelated to real words in English (O-P- controls). A developmental study by Goswami et al. (2001) demonstrated that for these same stimulus items, native German-speaking children showed more of a disadvantage in rejecting pseudohomophones than native English-speaking children. Unlike the Goswami et al. study, adult German-English bilinguals in Experiment 1 did not show a larger pseudohomophone disadvantage in English than the native English speakers. In addition, there were no overall differences in syllable length effects for native speakers of German and native speakers of English. Because German is more consistent than English, if native German speakers were using L1 processing strategies on this task, a larger syllable effect should be revealed. The Japanese-English bilinguals in Experiment 1 did show a larger syllable effect than the other groups. Because the three alphabets of Japanese all have high spelling-to-sound consistency, this large syllable effect suggests that this native speaker group was relying somewhat on L1 processing strategies. However the similarity of their performance on nonword responses with the native English speaker group also suggests that they were able to adopt L2 processing strategies under some conditions in this task. A post-hoc analysis of proficiency for the Japanese-English group suggested that the degree to which the L2 strategy was adopted depended on proficiency. Surprisingly, the less proficient Japanese-English participants had shorter response latencies and greater nonword accuracy than the more proficient participants. This suggests that the higher proficiency group was performing a checking strategy on the nonwords in this study and that this checking strategy may have been driving the differences between this group and the native speakers of English and German. In contrast, proficiency did not appear to influence the pattern of data for the German-English bilinguals.

Overall, the most surprising finding in this study was that there were no significant differences in the pseudohomophone effect across the three language groups It appears that despite the greater consistency of German and Japanese than English these two non-native speaker groups appeared to be able to successfully use processing strategies in English that were similar to those used by native speakers. One reason for the observed similarities may be the nature of the materials. Because the Goswami et al. (2001) materials were designed to be used with children, the high frequency of the words may have facilitated English-like processing by non-native speakers.

In Experiment 2, the lexical decision task was again used to contrast the performance of native English speakers and German-English bilinguals. However the task used for this experiment was an adaptation of a clustering paradigm used by Treiman and Chafetz (1987) in which participants were presented with an entire string of letters that was divided by slashes either after the onset or after the first vowel. Furthermore, the materials had been designed originally to be used with adult native speakers of English. Despite these changes, the results for both language groups in Experiment 2 were largely similar to one another and to those of Experiment 1.

A surprising finding in Experiment 2 was that the two new variations of the clustering task modeled on Treiman and Chafetz's (1987) paradigm failed to replicate their results. Treiman and Chafetz found that native English speakers were faster to respond to words and nonwords that were divided by onset and rime, in line with the parsing that would be hypothesized to be preferred in English. In Experiment 2, contrary to the Treiman and Chafetz results, words and nonwords parsed after the first vowel had faster response times than onset-rime parsed stimuli. Although the results of Experiment 2 did not replicate the earlier study, the benefit for nononset parsed stimuli was consistent across the two methods of presentation and the three SOA conditions, suggesting that this task was indeed sensitive to parsing.

One possibility why our findings may have differed from those reported by Treiman and Chafetz (1987) is that reaction times to both words and nonwords were substantially shorter in the present experiment than in the Treiman and Chafetz study. The longer RTs in the previous study suggest that participants may have been using a late checking strategy before making a lexical decision. If they were checking the letter string before making a decision, the fact that the letter string was divided by slashes may have had the effect of making the rime more salient. In contrast, in Experiment 2, participants may have been more likely to use the vowel in the nononset parsed words to help narrow down the list of possible rimes that could follow. In this case, having the vowel provides extra information which allows processing to unfold more rapidly. It is also possible that participants did benefit more from onset-rime parsed words, but that benefit may have been obscured by the effect of having the extra information from the vowel. Because most current models of word recognition operate in cycles, it is difficult to determine which of the two possibilities mentioned above is correct.

The possibility that the Treiman and Chafetz (1987) results may reflect a strategic effect is problematic for cross-linguistic replications of this study. Martensen, Maris, and

Dijkstra (2003) used the same slash clustering method for examining parsing of native Dutch speakers in Dutch. They found that clustering Dutch words revealed processing disruptions slightly different than those found by Treiman and Chafetz (1987) and that these clustering patterns were consistent with sublexical units important in Dutch. However, the participants from Martensen et al. (2003) also had faster overall response times than the participants from the Treiman and Chafetz (1987) study making it difficult to determine whether the differences observed cross-linguistically are due to differences by the structures in Dutch and English or to strategy differences between the two participant groups.

As for the comparison between the native English and native German participants, Experiment 2 again revealed that lexical parsing for the two groups was generally similar. Both showed an advantage for nononset rime parsing, and both also had shorter response times in the longer SOA conditions. For native English speakers, the SOA condition at which processing was least disrupted was the 250 ms condition. For native German participants, the minimal disruption occurred at the 500 ms SOA. This finding for SOA suggests that participants could maximize their use of the presented information if it was presented at a particular interval. Since non-native speakers generally process L2 words more slowly than native speakers, the optimal processing window may have been displaced to a later point in time.

The findings of Experiment 2 ultimately did not replicate Treiman and Chafetz (1987) but the task did provide a useful method for examining parsing and its time course. Although there were few differences in processing strategies between the native and non-native groups, the data suggest that the unfolding of processing for the native

German participants was occurring at a slightly slower pace than the native English participants. This was expected however since the native German participants were performing lexical decision in their L2.

Collectively, the lexical decision results in Experiments 1 and 2 did not distinguish parsing preferences between native English and native German speakers. However, in lexical decision participants do not have to rely on phonology to make decisions about strings of letters. Phonology may be active but the precise phonology associated with a given word does not have to be specified since spoken responses are not required. Since German and English differ in the consistency of spelling-to-sound mappings, Experiment 3 compared the two groups again but in a naming task to determine whether the requirement to produce the phonology associated with a specific lexical entry would be more likely to reveal cross-language parsing preferences. If the requirement to retrieve the phonology that will allow articulation to take place imposes additional language-specific processing requirements, then we might be more likely to see differences between native and non-native speakers under these conditions.

In Experiment 3, participants named words and nonwords. The words were either cognates or words that were noncognates but highly similar phonologically to German words. These words and nonwords varied in length and also in number of body neighbors. If the native German speakers use L1 processing strategies when naming words in L2, they should show larger length effects and smaller body neighborhood effects than the native English speakers. The results of Experiment 3, like the results of Experiments 1 and 2, showed that native German participants generally demonstrated similar processing strategies as native speakers of English. There were no differences in the length and body neighborhood size effects for nonwords, items whose phonology is generally thought to be computed, suggesting that the German participants used L2 strategies during this task.

However, the native German speakers did differ significantly from the native English speakers in their performance in naming real words. For words, there was an interaction which revealed that body neighborhood size interacted with word length only for the native German participants. For these non-native speakers, the length effect was stronger for words with small numbers of body neighbors than for words with large numbers of body neighbors. There are a number of reasons why this may have occurred. First, words with large numbers of body neighbors have many other English words that are similar. This in turn should make these items seem more English-like to non-native speakers. In addition, the small body neighbor word condition had more identical cognates than the large body neighbor condition. Because identical cognates were most likely to activate L1 processing strategies, it is not surprising that the length effect for these items was more similar to the native German participants performing in German in Ziegler et al. (2001) than the native English participants in Experiment 3.

Perhaps what is most striking about the results of Experiment 3 is that the presence of many cognates did not generally encourage greater use of L1 processing strategies overall for the native German speakers. Rather, the pattern of performance appeared to revert to the L1 strategy under local conditions that were guided by the information that was activated by the stimulus word itself.

Lexical Parsing and Current Models of word recognition

In Chapters 1 and 2, two prominent models of word recognition were discussed, the BIA and BIA+ models of bilingual word recognition (Dijkstra & van Heuven, 1998, 2002) and the Dual Route Cascaded (DRC) model (Coltheart et al., 2001). In addition, in Chapter 2, three bilingual variations of the DRC were proposed, one in which the bilingual has two separate dual route systems, one in which L2 information is integrated within the direct route, and a third in which L2 information is integrated in both the direct and the assembled routes.

The BIA+ and DRC models operate under different assumptions and therefore make different predictions about the locus of parsing effects in each of the three tasks used. In the BIA+ model there are two systems, an identification system and a task schema system. In the identification system, a stimulus first activates sublexical orthography which in turn activates sublexical phonology and lexical orthography, and finally lexical phonology and semantics. Sublexical information spreads activation to the language nodes and other language competitors are suppressed. The task schema system includes all of the strategies associated with the task such as the criteria that needs to be reached for a decision to be made. Because the language nodes only contain tags for determining which words belong to which language, the task schema system seems to be the most likely locus for lexical parsing strategies. Having these strategies in the task schema system places them outside of the lexicon and suggests that lexical parsing occurs relatively late in processing, after cross-language interactions have already occurred. The findings from Experiments 1 and 2 are consistent with this model because the data from these lexical decision experiments can potentially be interpreted as occurring relatively late in processing. However the presence of the interaction between body N size and word length in Experiment 3 suggest that the locus of parsing is much earlier than the BIA+ model would predict. The observed interaction suggests that language-specific strategies influence the activation of each language and that when L1 cues are sufficiently salient, the L1 processing strategies overshadow L2 processing strategies.

In contrast to BIA+, the DRC model embeds lexical parsing routines within the lexicon in the assembled route. If there is parallel activation of the assembled and addressed routes, we might then assume that it would be possible to observe effects of parsing early in processing. According to the DRC model, the direct route contains lexical orthography and phonology and retrieval of this information occurs in parallel. In the assembled route, phonology is assembled via language-specific rules. In addition to placing the locus of effects early in processing, the DRC model also predicts that longer stimulus items will take longer to process than shorter items given the serial nature of the assembled route. However one problem that this model has is that there is currently no place within the model for the storage of language information. Therefore in the third bilingual version of this model discussed earlier, the relative activation of the strategies for one language or the other would be guided by how much activation there is in the direct route.

The three experiments reported here demonstrated that highly proficient German-English bilinguals were able to process words like native English speakers in lexical decision and in naming when the data-driven aspects of stimulus processing did not specifically activate lexical candidates in L1. In addition, the interaction of word length and body neighborhood size in Experiment 3 suggests that the presence of many German-like words influenced parsing and that this effect occurred relatively early.

The DRC model seems able to handle both the effects of early parsing and also the length effect for syllables in Experiment 1, for phonemes in Experiment 2, and for letters in Experiment 3. Now the question that remains is which version of the bilingual DRC model best captures the findings across all three experiments.

Model 1 (refer to Figure 2.1) contains two independent dual route models and assumes that lexical access is language selective and that the bilingual can turn off one language while using another. Although processing for the proficient German-English bilinguals was extremely similar to the native English speakers, the evidence for L1 strategies in Experiment 3 suggests that this model does not adequately represent bilingual language processing. Furthermore it fails to account for the body of evidence on language nonselectivity in word recognition reviewed earlier.

The second proposed bilingual DRC model (refer to Figure 2.2) assumes that non-native speakers can only acquire the orthography and phonology of L2 words through the direct route. This lexical model resembles the procedural/declarative model proposed by Ullman (2001) for the lexicon and grammar for L2 which suggests that late L2 learners are able to acquire declarative but not procedural knowledge associated with the L2. According to the model, with increasing age of acquisition the ability to access procedural information is reduced. To the extent that procedural memory underlies the use of rules (in Ullman’s model, grammatical rules), late L2 learners will be required to memorize forms in the L2 that otherwise might be computed in the L1. Extending this to the lexical level suggests that L2 speakers can acquire information associated with the direct route, but only L1 information can take advantage of the assembled route. In Experiment 3, this model predicts the benefit for body N size that is similar between L1 and L2 speakers, however because only L1 strategies exist in the assembled route, this model fails to predict the overlap between native English and German speakers for nonwords in Experiment 3 and also the length similarities for these two groups across the three experiments. However, it is possible that because both German and English are orthographically deep languages, native German speakers were able to use existing L1 strategies in a way that emulated native English speakers. To determine if this is the case, future experiments will be needed to compare L2 strategies of speakers whose L1 varies in orthographic depth.

The model that best describes the findings reported here is a bilingual dual route model in which L2 information is integrated both within the direct and assembled routes (refer to Figure 2.3). In this model, L2 information is present but weaker in both routes of the model. Therefore in general, L2 processing strategies are used to process L2 words, however if L1 cues are present, or if the L2 processing strategies are not sufficiently strong, L1 processing strategies will be used during L2 processing. This model is able to describe both the body effect similarities and the length effect similarities across the three experiments. The locus of the body neighborhood effects in this model most likely occur in the direct route. As L2 speakers acquire more and more words that share a particular word body, the phonology for the word body becomes stronger and therefore easier to produce. Proficient bilinguals are likely to have a sufficient number of words from each body neighborhood to reveal an effect that closely resembles the one observed for native speakers. Both the third version of the bilingual DRC model and the BIA+ seem able to handle these results. However, because the locus of strategic information in BIA+ is much later in the stream of processing, BIA+ has more difficulty explaining the effects of length found in this experiment and for the nonword processing for Experiment 1. The length effects found in these two experiments appear to have a serial component and are most likely best captured by the assembled route of the proposed bilingual DRC. The nonwords in Experiment 1 were designed to examine strategies for the mapping of spelling-to-sound. The three speaker groups in Experiment 1 all demonstrated similar patterns for the three types of nonwords included in the study (O-P-, O+P+, and PSHs). This similarity suggests that within the assembled route for the two non-native speaker groups, L2 processing strategies were active and able to guide word recognition.

In addition, the third bilingual DRC model can account for some of the group differences found in the three experiments. In Experiment 1, although native Japanese speakers showed similar effects for responses to the three types of nonwords, they also had stronger effects of syllable length than the native English and German participants. Although this finding may have been due to their slower overall reaction times, it also suggests that some L2 processing strategies had not yet been acquired or were too weak to compete with L1 processing preferences. In Experiment 3, there was a much clearer case in which L1 processing strategies overcame L2 processing preferences. In this experiment, German-English bilinguals tended to process small body N words more like German words than English words. For these items, there was a stronger effect of length than for words with large numbers of body neighbors. The presence of a larger number of identical cognates in this condition may have encouraged L1 processing strategies for these items. It is for these effects that the models diverge. The third and final bilingual dual route model with L1 and L2 processing strategies in the assembled route can account for the early effects of parsing revealed by the body by length interaction in Experiment 3. In contrast, the BIA+ model would predict stronger effects of L1 processing strategies since the presence of cognates would maintain the activity of German throughout the stimulus list.

Conclusions

The results of the three experiments reported suggest that proficient non-native speakers are able to acquire processing strategies that resemble those of native speakers. For native German speakers in particular, it was clear that L1 processing strategies were not used because the pattern of results was more similar to native English speakers than to native German speakers performing either the same or a highly similar task in German. This performance difference was especially clear in the comparison of the native German participants in Experiment 3 and the native German participants in the Ziegler et al. (2001) study. In the present study and in Ziegler et al., the native German speakers were university students and likely to have had similar levels of proficiency in English. Both groups also performed the experiment in their native country. However despite the similarities between the two groups, the group performing the task in English appeared to be processing the cognates more similarly to native English speakers than the native German participants who named similar items in German. The pattern of results from these experiments also suggests that the ultimate attainment of lexical parsing strategies may differ depending on the L1. From Experiment 1, it is clear that the native English and native German speakers processed words and nonwords more similarly than native

Japanese speakers, although the native Japanese speakers overall did not differ from the native German speakers in rated proficiency and, if anything, by virtue of the fact that they were currently studying in the United States, had more exposure to English on a daily basis than the German participants. Although it is still possible that the differences in processing preferences may have been due to differences in proficiency, there is also the possibility that these differences arose because Japanese is more distinct from English in both orthography and consistency.

To distinguish whether the differences across the three groups for Experiment 1 occurred because of differences in proficiency or differences in L1 similarity, further research is needed with groups that are both more variable in proficiency and also in how similar the first language is to English. In addition, the results from Experiment 3 demonstrate an interaction between language selectivity and processing preferences that has not been previously explored. To further examine this interaction, it will be necessary to examine processing preferences over conditions that vary over a greater scale in how many L1 cues are present. Through further experiments and development of a bilingual dual route model, a greater understanding of how these two factors work together during bilingual language processing may emerge.

This further research should help resolve the paradox of how bilingual research demonstrates that both languages of a bilingual are active while cross-linguistic research reveals unique processing preferences across different languages. The BIA+ model suggests that during language processing, universal sublexical units are activated in a bottom up manner and that later in processing, language specific information determines which language is ultimately selected. This model seems to be able to capture a system where both languages can be active and competing but where language specific strategies guide selection late in processing. In contrast, in the third version of the bilingual DRC model proposed in Chapter 2, language-specific strategies guide processing early. The early influence of language-specific strategies is consistent with results from Experiment 3 and the interaction found for length and body N size found for only the native German speakers. However from the results of Experiment 3, it is unclear how early in the processing stream language-specific strategies take effect. In the monolingual version of the dual route model, the assembled and the direct routes begin processing at the same time. Given the assumption of parallel activation of the two routes, it is unclear whether in the bilingual case, the interaction for the native German speakers would be revealed. This version of the model generally predicts that the native German speakers and the native English speakers should perform similarly for all stimuli. However if the assembled route was offset relative to the direct route so that processing begins after some activation of items in the direct route has occurred, it may be possible to differentially engage language-specific strategies as a function of the nature of the activated information. In this adapted version of the final model, at the earliest stages the orthographic lexicons from both languages would become active. Language-specific processing strategies would activate depending upon which language was more active at a particular point. In Experiment 3, the low body N cognates most likely provided more activation to the German lexicon than to the English lexicon allowing the native German processing strategies to become active. In contrast the high body N words were likely to activate many other words in the English orthographic lexicon, allowing L2 processing strategies to become more active than L1 processing strategies.

In some respects the interaction found in Experiment 3 is problematic for both the BIA+ model and also the third version of the bilingual DRC model proposed in Chapter 2. For BIA+, the evidence for the early influence of processing strategies is at odds with the model's placement of these strategies relatively late in the stream of processing. The evidence for the influence of processing strategies early in processing is more closely aligned with the bilingual version of the dual route model. However it is not apparent how the DRC model decides which language processing strategies to use and whether these strategies begin unfolding immediately as predicted by the architecture of the model. These findings leave the paradox presented earlier unresolved. One possible resolution of this paradox would be to assume that the native German participants were using L1 processing strategies but just weighting these strategies differently because they were performing an L2 task. If so, then the second version of the bilingual dual route model with only L1 strategies in the assembled route and the BIA+ model both make similar predictions and could handle the results of the three experiments. It is possible that because English and German are both orthographically deep languages, German-English bilinguals are able to manipulate L1 strategies to make them appear like L2 strategies. To determine whether this is the case, it will be necessary to look at English processing in non-native speakers whose first language is orthographically shallow and therefore does not have the multiple letter consistency rules that are found in both English and German. If proficient bilinguals with an orthographically shallow first language are unable to demonstrate English processing preferences, then it suggests that what occurred in these experiments was a shifting of L1 strategy weights instead of the acquisition of L2 strategies. However if these bilingual groups are also able to consistently demonstrate L2 processing preferences, this suggests that the bilingual has acquired L2 processing preferences and the paradox remains unresolved.

In future research, it may be possible to determine when these parsing preferences unfold and to distinguish the two models by using methods that are more sensitive to the time course of processing. For example, ERPs may be able to distinguish between processing an unnaturally parsed word and a word that is more naturally parsed. Only by looking at a number of languages that vary in orthographic depth over conditions that are sensitive to the time course of processing can we determine which processing preferences are active and how they unfold.

REFERENCES

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. Psychonomic Bulletin & Review, 4, 439-461.

Bates, E. & MacWhinney, B. (1989). Functionalism and the Competition Model. In B. MacWhinney & E. Bates (Eds.), The crosslinguistic study of sentence processing (pp. 3-73). New York: New York Academy of Sciences.

Carreiras, M., Alvarez, C., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. Journal of Memory and Language, 32, 766-780.

Carreiras, M., Perea, M., & Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. Journal of Experimental Psychology: Learning, Memory, & Cognition, 23, 857-871.

Coltheart, M., Davelaar, E., Jonasson, J., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), Attention and performance VI (pp. 535-555). London: Academic Press.

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, 204-256.

Cuetos, F., & Mitchell, D. (1988). Cross-linguistic differences in parsing: Restrictions on the use of Late Closure strategy in Spanish. Cognition, 30, 73-105.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1989). Limits on bilingualism. Nature, 340, 229-230.

De Groot, A., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. Language Learning, 50, 1-56.

Dijkstra, T., & van Heuven, W. (2002). The architecture of the bilingual word recognition: From identification to decision. Bilingualism: Language and Cognition, 5, 175-197.

Dijkstra, T., & van Heuven, W.J.B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A.M. Jacobs. (Eds.), Localist Connectionist Approaches to Human Cognition. Mahwah, NJ: Lawrence Erlbaum Associates.

Dijkstra, T., van Heuven, W., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. Psychologica Belgica, 38, 177-196.

Dijkstra, T., van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. Bilingualism: Language and Cognition, 1, 51-66.

Dussias, P. (2001). Bilingual sentence parsing. In J. L. Nicol (Ed.), One Mind, two languages: Bilingual language processing (pp. 159-176). Cambridge, MA: Blackwell Publishers.

Frazier, L. (1979). On comprehending sentences: Syntactic parsing strategies. University of Connecticut doctoral dissertation.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), Attention and Performance XII: The Psychology of Reading (pp. 559-586). Hillsdale, NJ: Lawrence Erlbaum.

Goswami, U., Ziegler, J., Dalton, L., & Schneider, W. (2001). Pseudohomophone Effects and Phonological Recoding Procedures in Reading Development in English and German. Journal of Memory and Language, 45, 648-664.

Harm, M. & Seidenberg, M. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. Psychological Review, 106, 491-528.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? Bilingualism: Language and Cognition, 1, 213-229.

Implementation of the French DRC. (n.d.). Retrieved July 11, 2003, from http://www.maccs.mq.ed.au/~max/DRC/FrenchDRC.doc

Jared, D., & Kroll, J. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? Journal of Memory and Language, 44, 2-31.

Jared, D., Levy, B., & Rayner, K. (1999). The role of phonology in the activation of word meanings during reading: Evidence from proofreading and eye movements. Journal of Experimental Psychology: General, 128, 219-264.

Johnson, N. F., & Pugh, K. R. (1994). Models of visual word recognition. Cognitive Psychology, 26, 240-346.

Katz, L. & Feldman, L. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 157-166.

Kessler, B., & Treiman, R. (1997). Syllable Structure and the Distribution of Phonemes in English Syllables. Journal of Memory and Language, 37, 295-311.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. Journal of Memory and Language, 33, 149-174.

Lotto, L., & de Groot, A. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. Language Learning, 48, 31-69.

MacWhinney, B. (1997). Second Language Acquisition and the Competition Model. In A.M.B. De Groot and J.F. Kroll (Eds.), Tutorials in Bilingualism: Psycholinguistic Perspectives (pp. 113-142). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Martensen, H., Maris, E., & Dijkstra, T. (2000). When does inconsistency hurt? On the relation between phonological consistency effects and the reliability of sublexical units. Memory and Cognition, 28, 648-656.

McClelland, J., & Rumelhart, D. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. Psychological Review, 88, 375-407.

Michael, E., & Gollan, T. (in press). Being and becoming bilingual: Individual differences and consequences for language production. In J.F. Kroll & A.M.B. De Groot (Eds.). Handbook of Bilingualism: Psycholinguistic Approaches. New York: Oxford University Press.

Mitchell, D., & Cuetos, F. (1991). The origins of parsing strategies. Conference proceedings: Current issues in natural language processing (pp. 1-12). University of Texas at Austin, Austin, TX.

Pallier, C., Colomé, A., & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access. Psychological Science, 12, 445-449.

Plaut, D., McClelland, J., Seidenberg, M., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. Psychological Review, 103, 56-115.

Rubenstein, H., Lewis, S., & Rubenstein, M. (1971). Evidence for phonemic recoding in visual word recognition. Journal of Verbal Learning and Verbal Behavior, 10, 645-657.

Scarborough, D., Cortese, C., & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. Journal of Experimental Psychology: Human Perception and Performance, 7, 3-12.

Schwartz, A., Kroll, J. F., & Diaz, M. (2001). Mapping orthography to phonology: More evidence for nonselective cross-language activation. Poster presented at the 42nd Annual Meeting of the Psychonomic Society, Orlando, FL.

Seidenberg, M., Petersen, A., MacDonald, M., & Plaut, D. (1996). Pseudohomophone effects and models of word recognition. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 48-62.

Taft, M., & Russell, B. (1992). Pseudohomophone naming and the word frequency effect. Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 45A, 51-71.

Treiman, R., & Chafetz, J. (1987). Are there onset- and rime-like units in printed words? In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 281-298). Hillsdale: Lawrence Erlbaum Associates, Inc.

Treiman, R., Mullennix, J., Bijelac-Babic, R., & Richmond-Welty, E. (1995). The special role of rimes in description, use, and acquisition of English orthography. Journal of Experimental Psychology: General, 124, 107-136.

Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. Bilingualism: Language and Cognition, 4, 105-122.

Van Hell, J., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. Psychonomic Bulletin and Review, 9, 780-789.

Van Heuven, W.J.B., Dijkstra, T. & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. Journal of Memory and Language, 39, 458-483.

Van Heuven, W.J.B. (2000). Visual word recognition in monolingual and bilingual readers: Experiments and computational modeling. Doctoral Thesis University of Nijmegen. NICI Technical Report 20-01.

Ziegler, J., & Perry, C. (1998). No more problems in Coltheart's neighborhood: resolving neighborhood conflicts in the lexical decision task. Cognition, 68, B53-B62.

Ziegler, J., Perry, C., & Coltheart, M. (2000). The DRC model of visual word recognition and reading aloud: An extension to German. European Journal of Cognitive Psychology, 12, 413-430.

Ziegler, J., Perry, C., Jacobs, A., & Braun, M. (2001). Identical words are read differently in different languages. Psychological Science, 12, 379-384.

Appendix A

Word and Nonword Stimuli for Experiment 1

	Nonwords				Words	
Syllable condition	Base Word (did not appear)	PSH	O+P+ control	O-P- control		
1 syllable	dull	dul	rull	faish	ball	knee
	fake	faik	dake	koog	band	lamb
	girl	gerl	rirl	cilf	beer	loud
	muff	muf	guff	foaj	blue	luck
	rage	raij	tage	soaf	boat	mouse
	robe	roab	tobe	veeb	chin	rice
	tape	taip	fape	zoash	film	rich
	turn	tirn	murn	ziop	foot	sand
					form	ship
					green	wine
					ground	wish
					hair	zoo
2 syllable	butter	*buttar	tutter	larvol	angel	honey
	coffee	koffi	loffee	hixa	answer	hotel
	comic	commick	womic	dantol	canoe	hunger
	little	littel	kittel	verpil	cellar	mother
	pillow	pilloe	tillow	ricop	clinic	music
	taxi	tacksi	paxi	terket	coffee	number
	ticket	tikket	bicket	perlem	concert	paper
	window	windo	tindow	sulrin	costume	pepper
					finger	sister
					football	summer
					garden	tiger
					hammer	uncle
3 syllable	banana	*bannanah	danana	lagido	Africa	magazine
	daffodil	daffoddyl	baffodil	dilotaf	argument	medicine
	dinosaur	dynosor	hinosaur	saurnosi	bakery	museum
	factory	facktori	dactory	rytorac	camera	positive
	hospital	hosspital	pospital	talpiros	December	president
	potato	pottatoe	fotato	kinterat	elephant	professor
	pyjamas	*pijamas	tyjamas	majanys	energy	radio
	tomato	*tomaytoe	pomato	todamo	everyone	restaurant
					family	serious
					grandfather	telephone
					industry	theatre
					instrument	wonderful

*Items changed to American English. The original items were butta, bannanarnar, pijarmas, and tomartoe respectively.

Appendix B

*Word and Nonword Stimuli for Experiment 2**

	Nonwords		Words	
Number of Phonemes	Onset break	Nononset break	Onset break	Nononset break
3 phonemes	CH***	CHA**	CH***	GNA**
	CHACK	CHAMB	CHECK	GNASH
	GN***	KNU**	CH***	SHA**
	GNICK	KNUCK	CHICK	SHACK
	TH***	SHI**	KN***	SHO**
	THACK	SHING	KNOCK	SHOCK
	TH***	THA**	TH***	THI**
	THAMB	THASH	THICK	THING
	WR***	WHE**	WH***	THU**
	WRESH	WHECK	WHICH	THUMB
	WR***	WRU**	WR***	WRE**
	WRISH	WRUCK	WRONG	WRECK
5 phonemes	BL***	BLI**	BL***	BLA**
	BLISP	BLIFT	BLAST	BLANK
	BL***	BRU**	BR***	BLI**
	BLUST	BRUST	BRISK	BLIMP
	CR***	CRI**	CR***	CRE**
	CRAST	CRIST	CRANK	CREST
	CR***	FLU**	CR***	CRU**
	CRIFT	FLUNT	CRISP	CRUST
	CR***	FRU**	DR***	FLU**
	CRISK	FRUNK	DRIFT	FLUNK
	DR***	GLA**	DW***	SCA**
	DREPT	GLARF	DWARF	SCALP
	DW***	GRO**	FR***	SKU**
	DWEST	GROSK	FROST	SKUNK
	PR***	SCO**	GR***	SPO**
	PRUNK	SCORF	GRUNT	SPORT
	SC***	SPI**	PR***	SWE**
	SCANK	SPINT	PRINT	SWEPT
	SK***	SWU**	SC***	SWI**
	SKOMP	SWUNK	SCARF	SWIFT
	SW***	TRA**	SL***	TRU**
	SWIMP	TRALP	SLEPT	TRUST
	TW***	TRE**	TR***	TWI**
	TWANK	TRELF	TRUNK	TWIST

*This list represents only one possible presentation of these items. See Chapter 4 for a description of all possible presentations.

Appendix C

Word and Nonword Stimuli for Experiment 3

	Nonwords		Words	
Length	Small Body N	Large Body N	Small Body N	Large Body N
3 letters	FAS	BRY	ACT	BAR
	FOO	DEE	ARM	COW
	HEG	FOT	BOX	DAY
	LER	GAT	BUS	FAN
	NAL	LAN	GAS	HAT
	NUP	LAT	PER	HAY
	REA	MOY	SEA	ICE
	SIL	NOP	TEA	PRO
	TIS	SAR	TOE	RAW
	TOF	SUT	ZOO	RED
4 letters	BOAM	BICK	BOAT	BANK
	FURK	GACK	FILM	BEER
	GOFT	LUNK	FLEA	BLUE
	NAUL	NIST	FORM	FOUR
	NOOF	NUCK	GOLF	MEAL
	PARN	PAMP	HAIR	NEST
	PERD	PLAR	LOUD	POST
	SAUN	TAIN	PLUS	RICE
	SIBE	TORD	TEXT	SAND
	TRUP	TUMP	VERB	WINE
5 letters	BROAL	BRIST	BEAST	BLOCK
	DROAN	BRUCK	CHAIR	DREAM
	GLIEF	CROST	CLOTH	NIGHT
	GLIRD	DRACE	CROSS	POUND
	MEAST	DRAIL	FRONT	SIGHT
	PLOAR	GROST	POINT	START
	PRISH	GRUCK	SCARF	STEEL
	SPILK	PLOCK	SPORT	STEEP
	SPOND	PROOM	STIFF	STICK
	STERK	STECK	STORM	STONE
6 letters	BAINT	DRIGHT	CHANGE	BOUGHT
	BLORCE	FLITCH	COURSE	BRIGHT
	BRALSE	FRATCH	FIERCE	FLIGHT
	FLURST	GRATCH	FREEZE	GROUND
	FREAST	PLENCH	LENGTH	PLEASE
	GLADGE	SCRAST	PHRASE	SCHOOL
	PLARCH	SPRAIL	PRINCE	SPRING
	STINCE	SPRAND	STRICT	STREAM
	STRAUL	SPRANK	TAUGHT	STRONG
	STROND	STROOM	WEIGHT	THRILL

[i] Special thanks to Eva Suarez for creating the German version of this task.

[ii] In the Ziegler et al. (2001) study some English-German word pairs were not cognates, only words that were similar in onset and word length. For example the word pair STORM-*STRAND* (meaning beach in German) was compared in the original study. Although not all words in this study were English-German cognates, the overwhelming majority were.

[iii] Two types of voice key errors that were found in this Experiment were either voice key failures in which the participant had to repeat the stimulus item to move to the next trial, or cases where the voice key responded to another sound other than the word onset (e.g., exaggerated breathing).

Bianca Moravec Sumutka
August, 2003

Education

2001-Present Ph.D., Psychology The Pennsylvania State University, University Park, PA

1998-2001 M.S., Psychology The Pennsylvania State University, University Park, PA

1994-1998 B.A., Psychology Loyola College, Baltimore, MD

Research Experience

1998 Spring – 1998 Summer National Institute on Drug Abuse Baltimore, MD
Student Intern/Research Assistant
Monique Ernst, MD, Ph.D. – Interviewed and tested children with ADHD using the Wisconsin Card Sorting Test, an N-Back memory task and other cognitive tasks. Assisted in researching for and writing papers.

1998 Fall – 2001 Spring The Pennsylvania State University University Park, PA
Research Assistant
Lael Schooler, Ph.D. – Worked to create an ACT-R model of lexical decision and conducted a study looking at neighborhood effects in English word recognition.

2000 Spring The Pennsylvania State University University Park, PA
Research Assistant
Deborah Kelemen, Ph.D. – Assisted with an adult study looking at objects and artifact categories. Tested children in a Functional Fixedness study.

2000 Summer Carnegie Mellon University Pittsburgh, PA
Summer Student
ACT-R Summer School and Workshop – Learned about the ACT-R Architecture and completed a small modeling project. Gave a student presentation during the workshop.

2002 Spring The Pennsylvania State University University Park, PA
Research Assistant
Judith Kroll, Ph.D. – Conducted a translation elicitation study to find cognate norms for Dutch words. In addition, I conducted a study looking at the differences in English word production between Dutch-English bilinguals and native speakers of English.

2002 Fall –2003 Spring The Pennsylvania State University University Park, PA
Research Assistant
Judith Kroll, Ph.D. – Conducted a study looking at the differences in English word production and recognition between native speakers of English and proficient non-native speakers of English. Assisted in the management of the Cognition and Language Laboratory.