EXPLORING THE NATURE AND IMPACT OF WORD ENDINGS AS ORTHOGRAPHIC CUES TO LEXICAL STRESS IN ENGLISH

by

Erin Sparks

Submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
April 2018

TABLE OF CONTENTS

LIST OF TABLES vii

LIST OF FIGURES viii

ABSTRACT x

LIST OF ABBREVIATIONS AND SYMBOLS USED xi

ACKNOWLEDGEMENTS xii

CHAPTER 1 1
Introduction

1.1 MULTISYLLABIC WORDS IN ENGLISH 2

1.2 LEXICAL STRESS IN ENGLISH 4

1.3 CUES TO STRESS IN WRITTEN WORDS 7

1.4 MODELS OF DISYLLABIC WORD READING 10

1.4.1 Dual-Route Model 11

1.4.2 Connectionist Model 12

1.4.3 CDP++ 13

1.4.4 Comparing Models of Disyllabic Word Reading 14

1.5 THEORETICAL ISSUES EXPLORED IN THIS DISSERTATION 16

1.5.1 Suffix and Non-Suffix Endings as Stress Cues 16

1.5.2 Mechanism Behind Readers' Learning of Stress Cues 20

1.5.3 Vowel Quality and Stress Placement 22

1.5.4 Stress Cues in Connected Text 25

1.6 SUMMARY OF DISSERTATION GOALS AND PREDICTIONS 28

CHAPTER 2 33
Suffix and Non-Suffix Word Endings as Cues to Stress in Written English

2.1 ABSTRACT 33

2.2 INTRODUCTION 33

2.2.1 Suffixes as Cues to Stress in English 34

2.2.2 Non-Suffix Endings as Cues to Stress in English 37

2.2.3 The Current Study 39

2.3 EXPERIMENT 1 42

2.3.1 Method 42
2.3.2 Results 47
2.4 EXPERIMENT 2 50
2.4.1 Method 50
2.4.2 Results 53
2.5 GENERAL DISCUSSION 55
2.5.1 Theoretical Contributions 58
2.5.2 Conclusions 60
2.6 APPENDIX A: EXPERIMENT 1 ITEMS 61
2.7 APPENDIX B: SUPPLEMENTAL FIGURES OF KEY STUDY FINDINGS 62
2.8 APPENDIX C: EXPERIMENT 2 ITEMS 65

CHAPTER 3 66
Getting to Know the Stress Neighbourhood: Statistical Learning of Word Endings as Stress Cues in English

3.1 ABSTRACT 66
3.2 INTRODUCTION 66
3.2.1 Word Endings as Cues to Lexical Stress 67
3.2.2 Statistical Learning of Stress Neighbourhoods 69
3.2.3 Differences in Readers' Exposure to Stress Neighbourhoods 72
3.2.4 The Components of Stress Neighbourhoods 75
3.2.5 The Current Study 77
3.3 METHOD 79
3.3.1 Participants 79
3.3.2 Materials 79
3.3.3 Procedure 83
3.4 RESULTS 84
3.4.1 Data Cleaning 84
3.4.2 Procedures for Model Construction and Comparison 85
3.4.3 Research Question 1: Effects of Proportion and Number of Stress Friends 89
3.4.4 Research Question 2: Effects of Number of Stress Friends and Number of Stress Enemies 95

3.5 DISCUSSION 99
3.5.1 Differences in Readers' Exposure to Stress Neighbourhoods 100
3.5.2 The Components of Stress Neighbourhoods 104
3.5.3 Theoretical Implications 105
3.5.4 Future Directions 106
3.6 APPENDIX D: LEXICAL DECISION TASK ITEMS 109
3.7 APPENDIX E: MODELS' RANDOM EFFECTS STRUCTURES 113

CHAPTER 4 115
Stress Guides Spelling Decisions Independently of Vowel Quality: Untangling the Gordian Knot
4.1 ABSTRACT 115
4.2 INTRODUCTION 115
4.2.1 Readers' Use of the Associations Between English Stress and Spelling 117
4.2.2 Theoretical Contributions of Understanding the Impact of Lexical Stress on Spelling Choice 121
4.2.3 The Current Study 122
4.3 EXPERIMENT 1: ADULT READERS 125
4.3.1 Method 125
4.3.2 Results and Discussion 132
4.4 EXPERIMENT 2: DEVELOPING READERS 136
4.4.1 Method 136
4.4.2 Results and Discussion 138
4.5 GENERAL DISCUSSION 142
4.5.1 Readers' Use of the Associations Between English Stress and Spelling 143
4.5.2 Lexical Stress and Models of Spelling Choice 146
4.5.3 Practical Implications 147
4.5.4 Limitations and Future Directions 148
4.6 APPENDIX F: PSEUDOWORD ENDING SPELLINGS 152

CHAPTER 5 .... **153**
Do Orthographic Cues to Stress Affect Reading in Connected Text?

5.1 ABSTRACT 153

5.2 INTRODUCTION 153

5.2.1 Ending Spellings as Lexical Stress Cues 154

5.2.2 Lexical Stress and Silent Reading 159

5.2.3 The Current Study 161

5.3 EXPERIMENT 1: ADULT READERS 164

5.3.1 Method 164

5.3.2 Results 169

5.4 EXPERIMENT 2: DEVELOPING READERS 173

5.4.1 Method 173

5.4.2 Results 175

5.5 GENERAL DISCUSSION 177

5.5.1 Ending Spellings as Stress Cues in Connected Texts 178

5.5.2 Extending Models of Isolated Multisyllabic Word Reading 183

5.5.3 Limitations and Future Directions 184

5.6 APPENDIX G: SUMMARY OF WORD ENDINGS' STRESS CONSISTENCY EFFECT SIZES 189

CHAPTER 6 .... **190**
Discussion

6.1 SUMMARY OF DISSERTATION GOALS 190

6.2 THEORETICAL CONTRIBUTIONS 191

6.2.1 Suffix and Non-Suffix Endings as Stress Cues 191

6.2.2 Mechanism Behind Readers' Learning of Stress Cues 193

6.2.3 Vowel Quality and Stress Placement 196

6.2.4 Stress Cues in Connected Text 198

6.2.5 Comparing Models on Dissertation Findings 200

6.3 AN ACCOUNT OF READERS' SENSITIVITY TO, AND USE OF, ENDINGS AS STRESS CUES ACROSS DEVELOPMENT 201

6.3.1 Lexical Stress in Oral Language Development 203

6.3.2 Early Reading Development 206

6.3.3 Later Reading Development 208
6.3.4 Skilled Adult Reading 212
6.4 LIMITATIONS AND FUTURE DIRECTIONS 214
6.4.1 Operationalizing Endings as Stress Cues 214
6.4.2 Individual Differences in Sensitivity to Stress Cues 217
6.5 CONCLUSIONS 222
6.6 APPENDIX H: ELEMENTARY GRADE FREQUENCIES OF WORD ENDINGS USED AS STRESS CUES IN RESEARCH 225

REFERENCES 229

LIST OF TABLES

Table 1.1. A comparison of three models of disyllabic word reading, summarizing the characteristics of each model that inform the theoretical issues explored in this dissertation. 32

Table 2.1. Descriptive statistics by condition, summarizing the lexical characteristics on which Experiment 1's conditions were balanced. 46

Table 2.2. Mean reaction times to correctly identified words and mean accuracy rates (with standard deviations) for Experiment 1. 48

Table 2.3. Descriptive statistics by condition, summarizing the lexical characteristics on which Experiment 2's conditions were balanced. 53

Table 2.4. Mean reaction times to correctly identified words and mean accuracy rates (with standard deviations) for Experiment 2. 54

Table 3.1. Zero-order correlations among all stress neighbourhood predictor variables and item-level control variables. 83

Table 3.2. Summary of the variance components (standard deviations) for each random-effects parameter in Models 1 to 4 (addressing research question #1). 113

Table 3.3. Summary of the variance components (standard deviations) for each random-effects parameter in Models 5 to 8 (addressing research question #2). 114

Table 4.1. Mean (and *SD*) PVI scores for pitch and intensity, quantifying the lexical stress contrasts in items pronounced with first-syllable stress and with second-syllable stress. 129

Table 4.2. Mean (and *SD*) formant frequencies of the vowels in pseudowords pronounced with first- and second-syllable stress. 130

Table 5.1. Summary of passage characteristics for the texts read by participants in Experiment 1 (adults) and Experiment 2 (children). 168

LIST OF FIGURES

Figure 2.1. Paired difference scores showing Experiment 1's effect of endings' stress consistency on lexical decision response times, presented as a function of ending type. 63

Figure 2.2. Paired difference scores showing Experiment 2's effects of endings' stress consistency (left) and ending type (right) on lexical decision response times. 64

Figure 3.1. Standardized fixed effects estimates for Models 1 through 4, predicting participants' log-transformed reaction times to correctly identified words. 91

Figure 3.2. Marginal effects [95% CI] of a word's proportion of stress friends and number of stress friends, respectively, as predictors of lexical decision latency in Model 4. 94

Figure 3.3. Standardized fixed effects estimates for Models 5 through 8, predicting participants' log-transformed reaction times to correctly identified words. 96

Figure 3.4. Marginal effects [95% CI] of a word's number of stress friends and number of stress enemies, respectively, as predictors of lexical decision latency in Model 8. 99

Figure 4.1. Mean rates at which participants chose the (A) simple spelling, (B) extended spelling, and (C) distractor spelling options as a function of age group and aural stress pattern. 133

Figure 4.2. Paired difference scores in the rate at which participants chose simple and extended spellings after items pronounced with first- versus second-syllable stress. 135

Figure 5.1. Mean rates at which participants omitted target letters as a function of ending condition. 170

Figure 5.2. Paired difference scores in the rates at which participants failed to detect target letters in words whose endings were consistent versus inconsistent with their stress patterns. 172

Figure 5.3. Forest plot showing the effects of ending cue consistency in published studies of English-speaking adults. 189

Figure 6.1. Boxplots and scatterplots showing grade-level frequencies of the word endings used in studies of English-speaking children's sensitivity to endings as stress cues. Frequency is quantified as the number of words in which each ending appears. 227

Figure 6.2. Boxplots and scatterplots showing grade-level frequencies of the word endings used in studies of English-speaking children's sensitivity to endings as stress cues. Frequency is quantified as the summed frequency of words in which each ending appears....................... 228

ABSTRACT

Lexical stress, which refers to the pattern of emphasis across syllables in a word, is central to the phonology and identity of multisyllabic words in English. As such, establishing how readers process stress in written words is crucial to a full understanding of English word reading. Recent evidence shows that written word endings act as probabilistic cues to stress, affecting readers' naming and recognition of written words (e.g., the ending *–et* is associated with first-syllable stress, whereas *–oon* is associated with second-syllable stress). In this dissertation, I address four open questions about the nature and impact of written word endings as cues to stress in English. Study 1 confirmed that English-speaking adults are quicker to make lexical decisions toward words whose endings accurately cue stress than those whose endings give misleading cues to stress. This effect was similar across endings that can and cannot serve as English suffixes, suggesting that suffixes are not unique in their role as stress cues. Study 2 explored statistical learning as a mechanism behind readers' sensitivity to word endings as stress cues. Its main finding was that adults' lexical decisions were affected by exposure to all words with a given ending—those that share a word's stress pattern facilitate quicker responses; those with differing stress patterns yield slower responses. Study 3 established that word endings cue lexical stress independently of vowel quality—stress' key phonemic correlate. In a spelling choice task that manipulated stress while controlling vowel pronunciations, adults and children in Grades 5–6 used aurally-presented stress patterns to guide their pseudoword spelling choices. However, younger children preferred the more frequent spelling option, regardless of stress pattern. Finally, Study 4 found that adults use word endings as cues to stress when reading texts for comprehension, though elementary school-aged children do not. Together, these findings speak to a robust role for word endings as orthographic cues to stress across a variety of reading tasks. I discuss implications for theoretical models of word reading and argue that the findings align with a connectionist account of stress placement in multisyllabic word reading.

LIST OF ABBREVIATIONS AND SYMBOLS USED

ANOVA	Analysis of variance
CDP++	Connectionist Dual Process model of disyllabic word reading
CELEX	"Centre for Lexical Information," a database of English words
CI	Confidence interval
d	Cohen's *d* estimate of effect size
dB	Decibels
ERP	Event-related potential
F	*F* test statistic for analysis of variance
F1; F2; F3	First, second, and third vowel formants, respectively
Hz	Hertz
kHz	Kilohertz
lme4	"Linear Mixed-Effects Models using 'Eigen' and S4," a statistics package used to conduct linear mixed-effects analyses in R
M	Mean
MANOVA	Multivariate analysis of variance
MLE	Missing letter effect
ms	Milliseconds
N	Total sample size
n	Subsample size
OLD	Orthographic Levenshtein distance
p	*p*-value indicating statistical probability
PCA	Principal component analysis
PVI	Pairwise variability index
RT	Reaction time
SD	Standard deviation
t	*t* test statistic for comparing means
U	Mann-Whitney *U* test statistic used for non-parametric comparisons of two samples
z	Standard score indicating the distance in standard deviations of a value from the mean
η_p^2	Partial eta-squared estimate of effect size
Λ	Wilks' lambda test statistic for multivariate analysis of variance
χ^2	Chi-square test statistic, used for likelihood-ratio tests

ACKNOWLEDGEMENTS

I might have started this dissertation for my own sake, but that's certainly not why I finished it. There are so many people who deserve more than just figurative credit for getting me to this point, and there are even more who made sure I got here with a smile on my face. I'm glad to have this chance to thank them for everything they've given me.

First and most notably, I'd like to thank my dissertation supervisor, Dr. Hélène Deacon. Hélène, you had more faith that I could do this than I ever did, and I'm so grateful for all the opportunities you've given me to grow as an academic and to take on new roles. More than that, thanks for helping me use those opportunities to build a career path that I can be proud of. Your expert guidance, support, and ongoing patience were instrumental to my success—you've given me a remarkable example of mentorship and academic rigor to strive for.

To the members of my dissertation committee, Dr. Ray Klein and Dr. Aaron Newman: Your feedback pushed me to clarify my thinking and explore new ideas. My dissertation benefitted tremendously from that guidance, and I have no doubt that it has made me a better researcher as well. I would also like to thank my external examiner, Dr. Linda Jarmulowicz, whose work I have admired since I was an undergraduate. I appreciate the critical eye you brought to this document and your thoughtful suggestions for how to build on its findings in future. It was an honour (and a real pleasure) to have you evaluate my dissertation.

It's hard to imagine what my time here would have been like without the wonderful Language and Literacy Lab (and I'm not sure I want to try!). So many lab members and alumni shaped my experience that I can't possibly mention everyone by name—but it's precisely that range of people and diversity of interests that made my time in the lab so valuable. Learning from each of you broadened my horizons and deepened my research skills. I'd glad to have a network of such smart and generous people, here in Halifax and throughout the world, who are my friends as much as they are my colleagues. Special thanks to Annie Laroche for keeping everything on the rails!

To my students: You reminded me why this degree was worth getting. Thanks for playing along in the beginning while I pretended to know what I was doing. Even bigger thanks for helping me reach a point where I (almost) believe that I do. I feel lucky to have played a small role in your education—you played an enormous role in mine.

The best advice I got before starting this degree was to maintain a life outside the university, and it's advice I'm glad I took. I'm grateful to my extended family, who always asked how my research was going, and my friends, who learned when not to ask. Special mention belongs to my grandmother, Linda Sparks, who's been giving me thesis defense tips for as long as I can remember.

Finally, to my family: Megan[*], thanks for distracting me with a running list of things-dogs-ate-but-shouldn't-have and the occasional picture of weird-looking spleens. Never forget that I beat you to the finish line by an hour (but congrats on your DVM anyway). Evan[†], thanks for refusing to call me Dr. Sparks. Feel free to keep that ego-check going long after my convocation. Mom and Dad, thank you for the lifetime of advantages that helped me get here and let me succeed. Your support of all kinds is what made this possible, and I'll always be grateful for that.

[*] …who wants it on record that she knows the word *rennet* (plus a wealth of other cow-related medical trivia). I'm not adding this footnote to page 7—sorry, kid.

[†] He probably knows about cow enzymes too, though it's less clear why.

CHAPTER 1
INTRODUCTION

Reading is inextricably linked with oral language, and the phonology of language is a notable source of that link. Years of research has explored the effects of phonology on reading in English, dealing primarily with the reading of monosyllabic words. Much of this work has focused on the role of phonemes—the individual speech segments that correspond with graphemes in alphabetic written language. Children's learning of grapheme–phoneme correspondences is central to influential theories of reading development (e.g., Ehri, 1995), and prominent computational models have been developed to represent the processes by which readers use grapheme–phoneme correspondences to read monosyllabic words (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Seidenberg & McClelland, 1989).

This line of research has been enormously valuable and productive, but its tight focus on monosyllabic words is subject to legitimate critiques (e.g., Perry, Ziegler, & Zorzi, 2010; Yap & Balota, 2009). Two important gaps drive many of those critiques. The first is that monosyllabic words are a minority of the English lexicon (Baayen, Piepenbrock, & van Rijn, 1995), and multisyllabic words present readers with unique challenges that do not apply to monosyllables (e.g., Yap & Balota, 2009). As such, a full understanding of word reading in English requires study of multisyllabic words. A second critique is that, in focusing on monosyllables, the field has largely neglected a key component of phonology in English—lexical stress (e.g., Wade-Woolley & Wood, 2006; Wang & Arciuli, 2015). Lexical stress refers to the pattern of emphasis across syllables in a word. In English, stress placement varies across words and is not marked explicitly in the writing system. However, recent evidence suggests that there are probabilistic cues to stress in written English words (Arciuli & Cupples, 2006; Arciuli, Monaghan, & Ševa, 2010; Kelly, Morris, & Verrekia, 1998; Monaghan, Arciuli, & Ševa, 2016). Specifically, the spellings of many word endings (e.g., *–ar*, *–et*, *–oon*, *–umb*) are statistically associated with particular stress patterns. Skilled and developing readers are sensitive to these associations, which affects their processing of multisyllabic written words (e.g., Arciuli & Cupples, 2006; Arciuli et al., 2010; Burani, Paizi, & Sulpizio, 2014; Kelly et al., 1998; Mundy & Carroll, 2013; Sulpizio, Arduino, Paizi, & Burani, 2013). However, questions remain about the exact nature, mechanism, and scope of word endings' role as stress cues in English. In this dissertation, I will examine several of these questions in order to better understand readers' processing of stress when reading multisyllabic English words.

1.1 MULTISYLLABIC WORDS IN ENGLISH

Despite researchers' longstanding focus on monosyllabic words, the vast majority of words in English have more than one syllable. According to corpus estimates, monosyllabic words make up only 10% to 15% of distinct English words (Arciuli et al., 2010; Monaghan et al., 2016), and the factors involved in reading this small subset of words may not generalize to the lexicon as a whole (Yap & Balota, 2009). This leads us to consider the factors involved in multisyllabic word reading. In the texts read by English-speaking adults, words with one to three syllables make up 82.4% of distinct words (and 96.8% of all word occurrences; Monaghan et al., 2016), capturing much of the lexicon. The same point is true of the texts read by young children, where multisyllabic words appear early and often. Instructional emphasis on reading multisyllabic words typically begins at around Grade 2 (Schwanenflugel & Knapp, 2015), and from that point forward the number of multisyllabic words that children encounter grows rapidly with each subsequent year of elementary schooling. To demonstrate this growth, Kearns and colleagues (2016) used a corpus of children's texts to estimate the total number of multisyllabic words that young readers encounter at each elementary grade level. Their analysis shows a particularly sharp increase in the occurrence of multisyllabic words between Grades 2 and 5. Within this range, each increase in grade level yields at least 19,000 additional encounters with a multisyllabic word (note that Kearns and colleagues' estimates capture the total number of multisyllabic word occurrences, not the number of unique multisyllabic words in children's texts). Notably, these multisyllabic words often carry the content-specific and meaning-laden information that allows readers to understand texts (e.g., Archer, Gleason, & Vachon, 2003; Bryant, Ugel, Thompson, & Hamff, 1999; Cunningham, 1998), making fluent and accurate multisyllabic word reading a helpful skill when reading for comprehension.

Although multisyllabic word reading draws on many of the processes involved in reading monosyllables (e.g., the application of grapheme–phoneme correspondences), it also presents a unique set of challenges. For example, it requires dividing a word into its component syllables, placing the appropriate boundaries in words like *car.pet* and *ca.reen* (Perry et al., 2010). As a further complication, the spelling–sound correspondences in multisyllabic words can be less predictable than those of monosyllables, with particular ambiguities in a word's vowels (Venezky, 1999). For instance, the spelling *–ain* is pronounced consistently across all monosyllabic words (e.g., *pain*, *main*, *rain*; Ziegler, Stone, & Jacobs, 1997), but its pronunciation varies in multisyllabic words (e.g., *contain* vs. *mountain*). A final challenge—and a crucial one, for the purposes of this dissertation—is that multisyllabic words have lexical stress patterns. Lexical stress is a key component of an English word's phonology and its lexical identity, and as such, successfully reading multisyllabic words involves placing stress appropriately.

1.2 LEXICAL STRESS IN ENGLISH

Lexical stress refers to the pattern of relative emphasis across syllables in a word. It captures a complex auditory percept whereby a combination of acoustic cues (duration, intensity, and pitch; Fry, 1958) leads to the impression that some syllables in a word are stronger and more prominent than others. Specifically, stressed syllables are longer, louder, and higher in pitch than unstressed syllables, and are consequently perceived as having stronger emphasis. English is considered a stress-timed language, which means that stressed and unstressed syllables alternate at a roughly constant rate (e.g., Ramus, Nespor, & Mehler, 1999). As a result, all multisyllabic words in English have a lexical stress pattern (Cutler, 2015). And despite the fact that lexical stress is a feature of oral language, it has recently gained traction among reading researchers for the role it plays in processing written language (collections by Thomson & Jarmulowicz, 2016, and Wang & Arciuli, 2015, provide diverse overviews of the field).

When considering the role of lexical stress in multisyllabic word reading, it is useful to make the distinction between segmental and suprasegmental phonology. Segmental phonology refers to individual speech sounds, or phonemes (the smallest meaningfully contrastive units of sound in a language). Traditional models of word reading heavily emphasize the links between segmental phonemes and written graphemes (e.g., Coltheart et al., 2001; Ehri, 1995), and decades of research has established that awareness of phonemes is a key contributor to reading development (e.g., Adams, 1990). Lexical stress, on the other hand, is a form of suprasegmental phonology—a feature that extends across individual speech segments (Hayes, 1995; Selkirk, 1980). Although there is some overlap between stress and phonemes in English (e.g., vowels in unstressed syllables often reduce to the phoneme schwa; /ə/), stress as a suprasegmental linguistic feature is largely separable from segmental phonology. When compared to the wealth of research into segmental phonology's role in reading, relatively little work has considered the potential role of suprasegmental phonology. This gap has become particularly apparent as the field moves toward a better empirical and theoretical understanding of reading multisyllabic written words (e.g., Kearns et al., 2016; Perry et al., 2010; Ševa, Monaghan, & Arciuli, 2009; Yap & Balota, 2009).

There are several ways in which lexical stress might be relevant to multisyllabic word reading in English. The link is perhaps most obvious when it comes to reading aloud—pronouncing a written word means that readers must produce a stress pattern. Yet stress may also be implicated in silent reading, from visual word recognition to reading of connected text. At the word level, stress in English is often lexically contrastive (but see Cutler, 2015); it distinguishes between two orthographically and phonemically similar words (e.g., the noun *REcord* vs. the verb *reCORD*[1]; *TRUSty* vs. *trusTEE*, etc.). In other cases, misplacing a word's stress pattern yields a pronounceable nonword (e.g., **cheMIST*; **POlite*), which impedes word recognition (Small, Simon, & Goldberg, 1988). As a result, identifying the correct stress pattern of written words may be an important part of lexical access when reading (with converging evidence to suggest that

[1] Throughout this dissertation, I will use capital letters to indicate stressed syllables.

stress specifically affects the late stages of lexical access; Ashby & Clifton, 2005; Gutiérrez-Palma & Palma-Reyes, 2008). Readers also seem to activate word-level stress when they are reading silently for comprehension. For example, when adults read sentences (e.g., "*Poverty is the most [fundamental / significant] problem in our society today*"), they spend more time fixated on words that have two stressed syllables (*FUNdaMENtal*) than words of matched length and similar frequency that have only one stressed syllable (*sigNIFicant*; Ashby & Clifton, 2005). They also use sentences to form expectations about the stress patterns of upcoming words (Breen & Clifton, 2011). This work has been interpreted in light of Fodor's (1998) implicit prosody hypothesis, which argues that readers impose stress and intonational information onto text as they read it. All told, the evidence to date suggests that readers do process stress when they read multisyllabic written words.

That said, placing stress in English words is not always a straightforward task. In some languages, stress is fixed; stressed syllables appear in the same position across words. For example, words in Finnish reliably have first-syllable stress (e.g., *KAHvi*, *HAArukka*), while words in Polish reliably have stress on their second-last syllable (e.g., *JUtro*, *seKUNda*). This leaves speakers and readers of those languages with clear rules to guide stress placement. In contrast, English is a free stress language—although the stress patterns of individual words are stable, stress placement varies across words (e.g., *RACKet* has first-syllable stress, whereas *racCOON* has second-syllable stress). Moreover, the English writing system does not mark stress directly—there are no diacritic markers to show readers explicitly where stress should be placed (cf. languages like Greek, where stress is marked explicitly; e.g., *γέρος* vs. *γερός*). Taken together, these features of English mean that when readers are faced with written words (for example, the words *rennet* and *alee*, which are likely unfamiliar to many), there are few clear-cut cues that they can use to determine stress placement. Fortunately, though, the English writing system offers certain probabilistic cues to stress placement. Exploring the nature and impact of these written cues to stress is the overarching goal of this dissertation.

1.3 CUES TO STRESS IN WRITTEN WORDS

Written word endings are one useful cue to lexical stress in disyllabic English words. Their role as stress cues derives from the fact that certain ending spellings tend to appear in words with first-syllable stress (e.g., *–et*, *–um, –y*), whereas other spellings tend to appear in words with second-syllable stress (e.g., *–ette*, *–umb, –ee*). These associations between word endings and lexical stress are probabilistic, emerging from statistical regularities in the written language (Arciuli & Cupples, 2006). Yet readers who are sensitive to the role that endings can play as stress cues may be able to leverage those regularities to guide their stress placement in written words. In the example I gave above, for instance, a reader who encounters the unfamiliar word *rennet* might recognize the ending sequence *–et*, helping them to correctly identify the word's first-syllable stress pattern. Likewise, recognizing the ending sequence *–ee* in *alee* might help them to determine that the word has second-syllable stress.

Kelly, Morris, and Verrekia (1998) were among the first to report on the associations between ending spellings and lexical stress in English (see also Verrekia, 1996). Specifically, they noticed that many of the English word endings associated with second-syllable stress have more letters than is strictly necessary to represent the ending's phonemes (that is, the endings contain silent letters and/or letter doublets; e.g., *–ette*).

Kelly and colleagues suggested that these additional letters might serve to orthographically mark words for second-syllable stress, helping to guide readers away from the first-syllable stress pattern that is most common in English words (e.g., Kelly & Bock, 1988). Interestingly, naming and visual lexical decision studies show that readers are sensitive to the written ending cues that Kelly and colleagues (1998) uncovered. Skilled adult readers are quicker and more accurate when responding to written words whose endings accurately cue their stress pattern (e.g., *COMet*; *rouLETTE*) than words whose endings give misleading cues to their stress pattern (e.g., *caDET*; *PALette*; Kelly et al., 1998). More recently, Mundy and Carroll (2013) have replicated these findings in skilled adult readers and reported the same pattern of results in adults with dyslexia. This early work was instrumental in demonstrating that readers are sensitive to the links between stress and spelling in English.

Although Kelly and colleagues' (1998) corpus work was certainly valuable, it had a notable limitation—the authors did not offer a precise definition of word endings, nor could such a definition be derived from the specific endings that they reported. For example, some of Kelly and colleagues' ending spellings represented a word's final phoneme (e.g., *–c* and *–que* in *basic* and *baroque*) whereas others were larger units (e.g., *–et* and *–ette* in *rivet* and *roulette*; Arciuli & Cupples, 2006). This flexible definition of word endings left researchers unable to use Kelly and colleagues' (1998) approach to explore the associations between stress and spelling systematically in English.

An influential dictionary analysis by Arciuli and Cupples (2006) sought to address this limitation. They proposed an operational definition of word endings that comprised "the letter string beginning at the second phonemic vowel of a disyllabic word (and including any following consonants)" (Arciuli & Cupples, 2006, p. 929). Effectively, this definition of word endings maps onto the rime of the word's final syllable. A key advantage of this definition is that it can be applied straightforwardly to nearly all disyllabic words in English, allowing for a comprehensive analysis of the extent to which written word endings are associated with stress. Arciuli and Cupples (2006) conducted such an analysis and confirmed that a wide range of word endings are associated with stress in English. The strength of these associations varies on a continuous scale. For example, 92% of words ending with *–et* had first-syllable stress, whereas that was only true of 62% of words ending with *–ew*; both endings are primarily associated with first syllable stress, but *–et* serves as a more reliable stress cue. Notably, Arciuli and Cupples' analysis showed that the associations between word endings and stress in English extend beyond the patterns that Kelly and colleagues (1998) observed. For example, the ending *–act* is strongly associated with second-syllable stress, though it is not "orthographically marked" with extra letters (per Kelly et al., 1998, p. 822). In the years since Arciuli and Cupples' (2006) dictionary analysis, similar corpus work has shown that word endings act as cues to stress in the texts read by young children (Arciuli et al., 2010), and that endings also cue stress in trisyllabic English words and across a variety of other stress-based languages (Monaghan et al., 2016).

The empirical work that I will present in this dissertation explores the role that word endings play as cues to stress in English. Throughout, I have adopted Arciuli and Cupples' (2006) definition of word endings—the rime of a disyllabic word's second syllable. While this approach has some drawbacks (which I will address in the General Discussion), it has several advantages that make it nicely suited to empirical research.

First, as noted above, the definition can be systematically applied to words in the English lexicon. This precision helps to limit the so-called researcher degrees of freedom that a more flexible definition of word endings might bring to the process of item selection (e.g., Wicherts et al., 2016). Next, Arciuli and Cupples' (2006) definition avoids issues related to ambiguous syllable boundaries—for instance, the ending for *raccoon* is *–oon*, regardless of how we syllabify the word (*ra.ccoon* vs. *rac.coon* vs. *racc.oon*). This feature of their ending definition is particularly attractive for work in English, as speakers of the language (e.g., Eddington, Treiman, & Elzinga, 2013) and dictionary sources (Marchand, Adsett, & Damper, 2009) often disagree on the syllabification of English words. Finally, the definition is linguistically motivated, mapping onto a psycholinguistic unit that is theoretically relevant to word reading (e.g., Ziegler & Goswami, 2005; cf. Monaghan et al., 2016, where several definitions of word endings were based on number of letters).

In sum, several corpus studies have shown that written word endings can act as cues to stress in English (Arciuli & Cupples, 2006; Arciuli et al., 2010; Kelly et al., 1998; Monaghan et al., 2016), and behavioural studies have shed light on how skilled and developing readers use those cues to place stress in written words (e.g., Arciuli et al., 2010; Kelly et al., 1998). This work is particularly relevant to our understanding of multisyllabic word reading in English and can be applied to the theoretical models that have been developed to address it.

1.4 MODELS OF DISYLLABIC WORD READING

For two decades, efforts have been made to expand traditional models of monosyllabic word reading to include disyllabic words (Perry et al., 2010; Rastle &

Coltheart, 2000; Ševa et al., 2009). This work reflects a qualitative step forward in modelling of English word reading, as it requires implementing mechanisms for translating print into stress patterns during word recognition and/or reading aloud. As a result, these models of disyllabic word reading facilitate our understanding of how readers process multisyllabic written words, and do so in a way that could plausibly scale from mono- and disyllabic words to the rest of the English lexicon. And notably for the purposes of this dissertation, these models also provide theoretical frameworks through which to explore the role of written word endings as cues to lexical stress.

Here, I discuss three models of disyllabic word reading, each of which takes a different theoretical approach to the challenges of stress assignment in English. Although several other models have been developed to include multisyllabic words (e.g., Ans, Carbonnel, & Valdois, 1998; Jouravlev & Lupker, 2015; Sibley, Kello, & Seidenberg, 2010), I focus on three that (a) have been implemented in English, (b) were designed with stress assignment as a primary focus, and (c) have been independently tested against skilled readers' performance when reading multisyllabic words (e.g., Mousikou, Sadat, Lucas, & Rastle, 2017).

1.4.1 DUAL-ROUTE MODEL

Rastle and Coltheart (2000) developed the first model aimed at assigning stress from a word's orthography. They adopted the structure of earlier dual-route cascaded models (e.g., Coltheart, Curtis, Atkins, & Haller, 1993), in which a word's pronunciation is determined through distinct lexical and sublexical pathways. The lexical route contains the spellings and pronunciations of known words, and is particularly useful when dealing with words whose pronunciation and/or stress pattern cannot be reliably established through rule-based procedures. In their model, Rastle and Coltheart (2000) did not implement a lexical route, though they proposed that stress could be included with entries in the phonological lexicon. Instead, their model focused on developing a working sublexical route, in which a word's pronunciation—including its stress pattern—is determined through a set of predefined rules. This process mirrors the rule-based application of grapheme–phoneme correspondences in traditional dual-route models of monosyllabic words (e.g., Coltheart et al., 1993). In Rastle and Coltheart's model, stress assignment rules were linguistically driven, based on the observation that affixes in English have fairly reliable effects on stress placement (Fudge, 1984). As such, they created the model's algorithm around a store of 54 prefixes and 101 suffixes, each of which either takes stress (e.g., *–ette*, as in *brunETTE*) or does not (e.g., *–ar*, as in *BEGGar*). Upon identifying a spelling that corresponds with one of these affixes (regardless of whether it functions as a meaningful affix in that word; e.g., *MORTar*), the algorithm uses the characteristics of that affix to place stress and determine vowel quality. When a word has no affixes, the model assigns first-syllable stress, reflecting the most common stress pattern found in English disyllables (Baayen et al., 1995). These rules for stress placement are coupled with a set of grapheme–phoneme correspondences and checks for phonotactic legality, allowing the model to generate both a stress pattern and a segmental pronunciation (Rastle & Coltheart, 2000).

1.4.2 CONNECTIONIST MODEL

In contrast to the dual-route approach taken by Rastle and Coltheart (2000), more recent work has used a distributed-connectionist framework to assign stress in disyllabic words. A model developed by Ševa, Monaghan, and Arciuli (2009) is a notable example.

Their goal in constructing this connectionist model of disyllabic word reading was to determine whether stress could be assigned to written words based solely on orthography, without a need for the elaborate rule-based system implemented by Rastle and Coltheart (2000). To that end, their model's architecture is fairly simple, consisting of slot-based letter units for orthographic input, 100 hidden units that learn associations between words' orthography and stress patterns, and a single output unit identifying stress placement. The model does not contain stores of affixes or lexical entries; instead, cues relevant to stress placement are learned through a training process designed to mimic readers' exposure to written words (Ševa et al., 2009). As a result, the model draws on statistical regularities in written English to assign stress, without a need to directly specify the relevant stress cues in advance.

1.4.3 CDP++

Finally, Perry, Ziegler, and Zorzi (2010) provide a well-developed model of disyllabic word reading in a recent iteration of their Connectionist Dual Process model (CDP++). As the name suggests, CDP++ takes a hybrid approach that combines key aspects of the connectionist and dual-route traditions. Like Rastle and Coltheart's (2000) proposed dual-route model, CDP++ has distinct lexical and sublexical pathways for translating print to sound. At the same time, the model's sublexical route consists of a two-layer associative network that functions much like a connectionist model (e.g., Ševa et al., 2009).

CDP++ incorporates stress in several places. First, its sublexical network contains two stress nodes, representing the associative network's predictions about stress placement. These sublexical nodes fully connect to the word's graphemes, and they learn to assign stress from spelling in much the same way that the model learns grapheme–phoneme correspondences (Perry et al., 2010). Thus, unlike Rastle and Coltheart's (2000) dual-route model of disyllabic words, the sublexical route in CDP++ does not rely on explicitly defined rules. The model's lexical route also stores information about stress—each lexical entry's pronunciation includes its stress pattern, which can be activated by retrieving whole words from the lexicon. Both routes connect to the model's stress output nodes, allowing a word's stress pattern to be generated from either source. This stress output subsequently combines with phoneme-level output to yield a complete word pronunciation. Unlike competing models (Rastle & Coltheart, 2000; Ševa et al., 2009), CDP++ is a fully-specified model that generates a stress pattern, segmental pronunciation, and produces reaction times in addition to output accuracy; this degree of completeness makes it a particularly useful model for comparisons against how readers process stress in written words.

1.4.4 COMPARING MODELS OF DISYLLABIC WORD READING

By building on the foundation provided by work on monosyllabic words, each of these models represents a step forward in our theoretical understanding of word reading. Research that compares their effectiveness at simulating human performance is ongoing, and shows that despite their differences, all three models are similarly effective at assigning stress to written words. For example, a recent study compared the three models on their stress placement of 915 disyllabic nonwords, and found that their responses aligned with the modal response from adult readers on 81% (CDP++), 79% (connectionist), and 73% (rule-based dual-route) of items (Mousikou et al., 2017). Despite their roughly comparable levels of accuracy, each model has areas of strength and weakness. For instance, CDP++ and the distributed-connectionist model are equipped to handle items with and without affixes, though both struggle at times to overcome a bias toward first-syllable stress when assigning stress to nonwords (Perry et al., 2010). Rastle and Coltheart's rule-based algorithm performs well on items with endings in its affix store, but does more poorly when faced with words that lack affixes (Ševa et al., 2009). Certainly, work that helps to further evaluate and refine the models' performance will be important in advancing our knowledge of the cognitive processes involved in reading multisyllabic words.

At the same time, several theoretical issues raised by these existing models of word reading have not yet been tested. Some arise directly from predictions made by the models, whereas others emerge when exploring gaps that exist in the models' coverage of stress assignment. Addressing these issues and gaps is a crucial step in ensuring that our models of multisyllabic word reading process lexical stress in ways that are cognitively plausible and that reflect the behaviour of skilled and developing readers. To that end, this dissertation explores four issues relating to readers' use of word endings as cues to stress. First, in Chapter 2, I consider the nature of word endings as stress cues, specifically evaluating whether suffixes play a special role as cues to stress. In Chapter 3, I test predictions that derive from a mechanism by which readers are thought to learn about endings' role as stress cues. Next, in Chapter 4, I address the link between vowels and stress in English, aiming to establish whether stress can affect readers' processing of written words independently of vowel quality. Finally, in Chapter 5, I explore the scope of word endings' ability to cue stress, testing whether they do so when reading words in connected texts. On some of these issues, the models make clearly opposing predictions, while on others there is some degree of consensus. In either case, the models' claims require empirical testing. Table 1.1 summarizes each model's approach to these four issues, and I elaborate on each in the sections that follow.

1.5 THEORETICAL ISSUES EXPLORED IN THIS DISSERTATION

1.5.1 SUFFIX AND NON-SUFFIX ENDINGS AS STRESS CUES

When examining the role of word endings as cues to stress, it is important to recognize that endings can be characterized in different ways. In an earlier section, I gave a definition of word endings that constitutes the rime of a word's second syllable, drawing from influential corpus work done by Arciuli & Cupples (2006). Under this definition, some word endings are purely orthographic units; their spelling patterns do not map onto permissible English suffixes (e.g., *–el* and *–uct*). On the other hand, many of these word endings are used as meaningful suffixes in English (e.g., *–ar*, as in *beggar*; *–ic*, as in *mythic*). However, it is not yet clear whether this distinction affects the extent to which endings act as cues to stress in English words. The existing models make contrasting predictions on this question. Rastle and Coltheart's (2000) model suggests that suffixes play a uniquely strong role as stress cues, whereas the connectionist and CDP++ models argue that suffix and non-suffix endings play equally strong roles (Perry et al., 2010; Ševa et al., 2009). As the first theoretical issue in this dissertation, I address this debate empirically. Specifically, I explore whether ending spellings that are used as English suffixes might serve as stronger cues to stress than do non-suffix endings.

There are certainly reasons to expect that suffixes might play a particularly strong role in stress assignment. Morphological theory makes clear-cut predictions about the stress-related properties of English suffixes—specifically regarding whether they shift or maintain the stress pattern of the base word to which they attach (Fudge, 1984). Importantly, research has established that readers are sensitive to this property of English suffixes. For example, when presented with nonwords, adults' judgements about stress placement align with the stress pattern predicted by the word's suffix (e.g., *bisTINity* is preferred over *BIStinity*; Wade-Woolley & Heggie, 2015). Developing readers show a similar sensitivity to the stress-related properties of English suffixes. When asked to apply a stress-shifting suffix to a provided base word (e.g., "Put *–ity* on the end of *active*"), the accuracy of children's stress placements improves steadily across the mid-childhood years, reaching 78% accuracy in 9-year-olds (Jarmulowicz, 2006). This behavioural work helps to affirm that suffixes are valuable cues to stress assignment in English.

At the same time, corpus analyses have shown that suffixes are not the only word endings to have reliable associations with stress in English. Indeed, some non-suffix endings have near-deterministic links with particular stress patterns; for example, 94% of disyllabic words ending in *–oon* have second-syllable stress (Baayen et al., 1995). In their corpus analysis of English word endings, Arciuli and Cupples (2006) noted that only 14% of the endings that they examined were suffixes. This supports the claim that non-suffix endings can serve as effective, albeit probabilistic, cues to stress. And as I will discuss in subsequent chapters, there is behavioural evidence to suggest that English-speaking adults (e.g., Kelly et al., 1998) and children (Arciuli et al., 2010) are sensitive to the links between non-suffix endings and lexical stress when processing written words.

The roles of suffix and non-suffix endings in stress placement is a clear point of divergence between the models of word reading described above. On the one hand, Rastle and Coltheart's (2000) dual-route model predicts a special role for suffixes. More specifically, it suggests a special role for "affix-like [letter] strings" (p. 350), which are word endings or beginnings whose spellings are those of an English affix, regardless of whether that spelling serves as an affix in a given word. By this definition, suffix-like endings that cue stress appear in words like *beggar* (where the ending *–ar* is a meaningful suffix), and also words like *mortar* (where it is not). This focus on suffix-like endings is made clear by the model's architecture. Specifically, its sublexical algorithm is built around a store of affix-like spellings (e.g., suffix-like word endings: *–ar*, *–ic*; prefix-like word beginnings: *dis–*, *re–*). These affixes serve as the model's only link between orthography and stress. Admittedly, the authors blur this line somewhat by including a small number of non-suffix endings (e.g., *–oo*) in their affix store, justifying the decision to do so by noting that these non-suffix endings share the stress-determining properties of suffixes (per Fudge, 1984). Even so, their model excludes the vast majority of non-suffix word endings, and their approach relies on the premise that suffixes have a strong—and largely specific—role to play in English stress assignment. Taken at face value, this argument suggests that suffix-like ending spellings (e.g., *–ar*) might have a stronger effect than non-suffix endings (e.g., *–el*) on the processing of stress during visual word recognition. Again, Rastle and Coltheart's (2000) model does not distinguish between real suffix uses of an ending (e.g., *beggar*) and pseudosuffix uses where the ending does not serve as a meaningful unit (e.g., *mortar*), suggesting that both suffix-like uses should have similarly strong effects on stress placement during word reading.

In direct contrast, the connectionist and CDP++ models place no special emphasis on suffix-like endings (Perry et al., 2010; Ševa et al., 2009). As Ševa and colleagues

(2009) argue, “such [morphological] cues are a part of the orthographic information, but, in the connectionist tradition, they do not have special status beyond other, equally reliable cues that may be present … to determine stress position” (p. 241). Both models allow their associative networks to learn of any relevant stress cues that are present in the orthography, and neither contains a dedicated list of affixes to which the learning algorithm can refer. Under this theoretical perspective, suffix-like endings and non-suffix endings (e.g., *–ar* and *–el*, respectively) should have similar effects on readers' processing of stress during word reading. The key factor is how strongly a given ending cues lexical stress, not the ending's morphological status.

Despite the theoretical importance of this issue, and the separate bodies of empirical evidence showing that suffix and non-suffix endings can each act as stress cues, no work to date has directly compared the two ending types. In Chapter 2, I address this gap with two experiments that test the models' contrasting predictions. In both experiments, adult participants completed a visual lexical decision task with items whose endings were either consistent with their stress pattern or inconsistent with their stress pattern. The first experiment compares performance on words with pseudosuffix endings (e.g., *MORTar*; *guiTAR*) to those with non-suffix endings (e.g., *CHAPel*; *hoTEL*), testing the claim that the former should have a particularly strong effect on stress placement in written words. The second experiment compares suffix and pseudosuffix uses of the same endings (e.g., *BEGGar* vs. *MORTar*), testing the premise behind Rastle and Coltheart's (2000) decision to not distinguish between the two suffix-like ending uses in their model. Taken together, these experiments offer key insight into the validity of assumptions that are central to existing models of disyllabic word reading.

1.5.2 MECHANISM BEHIND READERS' LEARNING OF STRESS CUES

Given the evidence that readers are sensitive to links between word endings and stress in English (e.g., Arciuli et al., 2010; Kelly et al., 1998), it is worth considering how they acquire that sensitivity. One frequently-invoked suggestion is that statistical learning might be the mechanism at hand—a possibility that has been put forward to explain empirical findings (e.g., Colombo, Deguchi, & Boureux, 2014; Sulpizio et al., 2013), and that has been implemented as a learning mechanism by computational models (Perry et al., 2010; Ševa et al., 2009). The central premise of statistical learning theory is that readers gradually and implicitly learn of regularities through exposure to input (Perruchet & Pacton, 2006)—in this case, through exposure to co-occurring spellings and stress patterns in English. That is, as readers encounter many words with the same ending cues (e.g., *COMet*, *JACKet*, *CARpet*; *caDET*), they come to recognize broader probabilistic regularities in those cues (e.g., the ending *–et* usually occurs in words with first-syllable stress). If statistical learning is indeed the mechanism by which readers learn of the associations between word endings and stress in English, several untested predictions can be made about their sensitivity to written stress cues. Testing some of these predictions is the second goal of this dissertation.

The connectionist and CDP++ models align quite well with the claim that statistical learning drives readers' sensitivity to word endings as stress cues. In fact, their associative networks are designed to instantiate statistical learning principles, detecting regularities in the orthographic input through exposure to a large corpus of English words. In the connectionist models designed by Ševa, Monaghan, and Arciuli (2009; see also Arciuli et al., 2010), the words presented during the model's training phase were sampled according to their frequency in English, mimicking the extent of readers' exposure to various words in texts. The sublexical route of CDP++ takes a slightly different approach to accomplishing the same goal, scaling the learning rate for trained words according to their written frequency (Perry et al., 2010). In both cases, the models are given more opportunities to learn frequently-occurring regularities than rare ones.

On the other hand, Rastle and Coltheart's (2000) dual-route model does not follow statistical learning principles. Instead, stress assignment via the sublexical route is based on an explicit set of predefined rules. Notably, the models' developers did not propose a mechanism by which readers might acquire those rules—this sets their model apart from its alternatives (Perry et al., 2010; Ševa et al., 2009), for which the underlying sublexical mechanism is central to the models' designs. That being said, Rastle and Coltheart (2000) do recognize that the associations between orthography and stress used by their algorithm comprise "statistically-based rules" (p. 345), and they acknowledge that a connectionist approach to stress assignment may be fruitful as a way to capture probabilistic regularities. This suggests that, although they did not specify an underlying mechanism for their model, they view statistical learning as a reasonable possibility.

In sum, all three models converge on the view that statistical learning is a plausible mechanism behind readers' use of word endings as stress cues in English. Despite this relative degree of consensus, there are key predictions of the statistical learning framework that have not been tested empirically. In Chapter 3, I aim to address this gap. English-speaking adult participants completed a visual lexical decision task, with items that varied in both the strength of their endings' associations with stress and the number of words demonstrating those associations. This design allowed us to test two predictions derived from the statistical learning framework: first, that readers should be most sensitive to stress cues that they encounter often, and second, that readers draw on both positive examples and negative counterexamples of a stress cue when processing stress in written words.

1.5.3 VOWEL QUALITY AND STRESS PLACEMENT

When shifting from monosyllabic to multisyllabic words, models of word reading are faced with the challenge of specifying whether and how stress should interact with segmental phonology. For models of English, this issue is particularly relevant to the link between stress and vowel quality. In English, vowels that appear in unstressed syllables are often reduced to the centralized, indistinct phoneme schwa (/ə/; for example, the first vowels in the words *canal* and *cocoon*). By contrast, stressed syllables cannot contain reduced vowels (e.g., Cutler, 2015; Ladefoged & Johnson, 2010). This characteristic of English results in a strong association between stress and vowel quality, connecting a word's lexical stress pattern to its segmental pronunciation in ways that make it difficult to isolate the effects of stress itself. In fact, a review of the oral language literature has led some to conclude that "English listeners do not attend much to suprasegmental cues in recognizing [spoken] words, but they do pay great attention to the pattern of strong and weak vowel realizations" (Cutler, 2015, p. 115). If true of written language, this claim presents a challenge for the empirical and theoretical work that examines the links between stress and spelling in English. To address that challenge, a third issue explored in this dissertation is whether readers use the associations between written word endings and lexical stress even when vowel quality—stress' key phonemic correlate—is controlled.

The existing models of disyllabic word reading each take a different approach to handling the link between stress and vowel quality. For instance, in the dual-route model proposed by Rastle and Coltheart (2000), stress and vowel quality are tightly connected. The model's algorithmic rules simultaneously determine a word's stress pattern and its segmental pronunciation, which means that decisions about vowel quality in a word co-occur with those of stress placement. Indeed, at many points in the Rastle and Coltheart (2000) model, the instructions to assign stress and reduce vowels appear as a single step. This interdependence means that the effects of stress and of vowel quality on word reading cannot be easily separated when empirical findings are applied to the model. And further, their model cannot rule out the possibility that readers assign stress based on vowel quality rather than relying on written cues to stress. As the authors themselves note, "in order to proceed with a serious model of disyllabic word reading, we would have to determine empirically whether readers derive cues for stress assignment from orthographic or from phonological [vowel quality] information" (Rastle & Coltheart, 2000, p. 357).

Taking the opposite approach, CDP++ assumes that stress and vowel quality are represented independently of each other (Perry et al., 2010). This is evident in the model's sublexical network, which contains stress nodes and phoneme nodes that both connect to a set of grapheme nodes, but that do not connect with each other (likewise, stress and phonemes are separated at the model's output level). Importantly for our purposes, this means that CDP++ learns of the associations between spelling and stress in ways that cannot be attributed to links between spelling and vowel quality. This is very much in line with the interpretation that written word endings can act as cues to lexical stress, and, by doing so, affect word reading in a way that extends beyond the role of segmental grapheme–phoneme correspondences (e.g., Kelly et al., 1998). Yet because of how tightly stress and vowel quality are confounded in English, further empirical work is needed to determine how accurately the CDP++ approach reflects reading behaviour.

In contrast with both alternative models, the connectionist model of disyllabic word reading avoids making assumptions about how stress and vowel quality interact (Ševa et al., 2009; see also Arciuli et al., 2010). In some ways, this open-ended approach is useful when dealing with an unresolved empirical question, as examining a model's performance might help to drive future hypothesis-building. For example, errors in which a connectionist model produces a schwa vowel in a stressed syllable (similar to errors found by Perry et al., 2010) would suggest that it has learned to place stress independently of vowel quality. Unfortunately, though, the current model's implementation is not complete enough to explore this link between stress and vowel quality. That is, the model's training input and word reading output are focused solely on stress placement; it excludes segmental phonology altogether. As a result, it is not clear if or how a more complete version of Ševa and colleagues' (2009) connectionist model would draw on vowel quality when processing lexical stress in written words.

These differing accounts of how stress might interact with vowel quality highlight the need for empirical research into the specific role of lexical stress when processing written English words. In Chapter 4, I address this issue by using a paradigm that manipulates the stress pattern of pseudowords while controlling vowel quality across pronunciations. In two experiments, adults and elementary school-aged children listened to pseudowords, then chose a preferred spelling from among forced-choice options

(Verrekia, 1996). Critically, the provided spelling options included endings that are phonemically similar, but that are associated with different stress patterns (e.g., the endings –*et* and –*ette*). By removing vowel quality as a confound, the experiments in Chapter 4 serve to establish whether readers are indeed sensitive to the associations between word endings and lexical stress in English.

1.5.4 STRESS CUES IN CONNECTED TEXT

All three of the models discussed in this chapter focus on stress placement when reading isolated words (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). As a result, the empirical research that tests and builds upon those models has also focused on words and pseudowords read in isolation (Arciuli & Cupples, 2006; Arciuli et al., 2010; Burani et al., 2014; Kelly et al., 1998; Sulpizio et al., 2013). This work has been useful in establishing readers' use of the links between stress and spelling in English. However, it is worth noting that isolated word reading is quite far removed from the everyday task of reading texts for comprehension. The two tasks place quite different demands on readers, and evidence suggests that measures of processing time for isolated word reading are only modestly related to those of reading text-embedded words (e.g., Kuperman, Drieghe, Keuleers, & Brysbaert, 2013). With this in mind, work on isolated word reading cannot tell us whether word endings act as cues to stress when they are read in connected texts. Testing the extent to which they do so is the fourth and final issue explored in this dissertation.

There is certainly evidence to suggest that written cues to stress are more useful to readers in some situations than in others. Even slight variations of the same task can affect endings' role in word reading. For example, a study of Italian readers found some evidence that word endings' association with stress affected lexical decisions—words whose endings were consistent with their stress pattern were recognized more quickly than words whose endings were not (Colombo & Sulpizio, 2015). Critically, though, this pattern only appeared in a high-demand version of the lexical decision task (where real words and foils had similar spellings, requiring in-depth processing). There was no such effect in a low demand version of the task (where words were more easily distinguished from foils; Colombo & Sulpizio, 2015). These discrepant findings suggest that we should not generalize our understanding of written stress cues across tasks—particularly not when tasks differ as dramatically as do isolated word reading and reading connected text.

In fact, a recent study suggests that even small amounts of context can influence readers' use of word endings as stress cues (Spinelli, Sulpizio, Primativo, & Burani, 2016). The study drew on the fact that, in Italian, word endings that are broadly associated with one stress pattern (e.g., *–ita* and penultimate stress) may be associated with a different stress pattern when they are used as grammatical markers (e.g., *–ita* as a third-person verb marker most commonly appears in words with antepenultimate stress). In this study, when readers encountered nonwords in isolation, they tended to produce the stress pattern most associated with the ending's spelling. However, presenting the same nonwords after a context word (e.g., a third-person pronoun) led readers to increasingly use the stress pattern linked to the grammatical properties of the ending. In other words, contextual information reduced the extent to which readers used the ending's spelling alone to guide their stress placement.

None of the existing models of disyllabic word reading are equipped to look at how stress is assigned to words read in connected text. Because these models focus on isolated words, they assign stress based solely on combinations of orthographic, phonemic, and/or lexically derived stress cues that are present in the word. In their current forms, the models cannot include any of the attenuating factors that may be present in connected text, and thus may inflate the role of ending cues in everyday reading for comprehension. Each model's authors acknowledge limitations in that regard, noting specifically that sources of semantic and/or grammatical information would help to improve their models' stress placement, and that if included, that information would likely interact with the written cues to stress (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). However, it remains to be seen whether the existing models will be able to account for the effect that surrounding text might have on the processing of stress in multisyllabic words (Spinelli et al., 2016). Research establishing the role of word endings in connected text is an informative step in that direction.

In Chapter 5, I explore the role of word endings as cues to stress when words are read in connected text. In it, I report two experiments—one with adults, and one with school-aged children—in which participants read a series of texts for comprehension. While reading, they were asked to mark a specified target letter each time that they noticed it. Crucially, these target letters appeared in words whose endings were either consistent or inconsistent with their stress patterns—a comparison that, to date, has only been explored using isolated word reading tasks (e.g., Kelly et al., 1998). The experiments in Chapter 5 extend this work to a paradigm involving connected text, which allows us to explore the scope of word endings' effect on word reading by testing whether they affect processing of stress during the everyday task of reading for comprehension.

1.6 SUMMARY OF DISSERTATION GOALS AND PREDICTIONS

In summary, the experiments that I will present in this dissertation explore the role of written word endings as cues to lexical stress in English. These experiments address four theoretically important issues by: (1) comparing the roles of suffix and non-suffix word endings as cues to stress [Chapter 2]; (2) testing predictions derived from a mechanism by which readers might learn of endings' roles as stress cues [Chapter 3]; (3) establishing whether readers' sensitivity to endings as stress cues exists independently of vowel quality [Chapter 4]; and (4) determining whether word endings act as cues to stress when read in connected text [Chapter 5]. Answering each of these questions will inform and extend the existing models of disyllabic word reading in English (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009).

All of the empirical chapters in this dissertation involve testing these four theoretical issues in skilled adult readers. In keeping with prior empirical work, I expected that, across tasks and manipulations, skilled adult readers would be sensitive to the links between stress and spelling in English (e.g., Arciuli & Cupples, 2006; Kelly et al., 1998; Mundy & Carroll, 2013). Chapters 2 and 3 involve visual lexical decision tasks, where I expected adults to respond more quickly to words whose endings accurately cue their stress patterns than words whose endings inaccurately cue stress. Beyond that broad prediction, the experiments' manipulations test my key theoretical questions. In Chapter 2, which compares suffix and non-suffix ending types, I expected that suffix-like endings might play a stronger role as stress cues than non-suffix endings (per the predictions of Rastle & Coltheart, 2000). In Chapter 3, which tests predictions of statistical learning theory, I expected that readers would be more sensitive to ending cues that appear in many words than those that appear in relatively few. Chapter 4 involves a different experimental paradigm, in which adults indicated their preferred spelling for an aurally presented pseudoword. On this task, I expected adults to preferentially choose spellings that are associated with the stress patterns of the pseudowords they heard. Importantly, this finding would show that sensitivity to the links between stress and spelling in English can exist independently of vowel quality. Finally, in Chapter 5, adults read texts for comprehension while detecting target letters in words whose endings were either consistent or inconsistent with their stress patterns. If word endings act as cues to stress when read in connected text, I expected to find a higher letter omission rate in words whose endings are consistent with their stress pattern. However, I suspected that this effect—if present—would be relatively small when compared with tasks that involve reading words in isolation.

None of the theoretical models reviewed above were designed to address developmental questions. The one exception is a developmentally-focused connectionist model created by Arciuli, Monaghan, and Ševa (2010). Conceptually and structurally, their model is nearly identical to the connectionist model that I described earlier (Ševa et al., 2009); thus, the same theoretical predictions apply to both models. The key distinction between the models is that Arciuli and colleagues (2010) aimed to simulate the process by which developing readers learn of links between stress and spelling in English. They did so by training their model on words from a corpus of children's texts (Zeno, Ivens, Millard, & Duvvuri, 1995), starting with words appropriate for 5-year-olds and progressing to words appropriate for 12-year-olds. Interestingly, the model showed the same developmental trajectory in the use of word endings to guide stress as did children of those same ages. That is, in a behavioural experiment, 5- to 8-year-old children made some use of word endings to guide stress, but tended to prefer first-syllable stress across word endings. By contrast, 9- to 12-year-old children showed a more robust use of word endings to guide stress placement (Arciuli et al., 2010). Taken together, the empirical and theoretical work done by Arciuli and colleagues (2010) is an important step toward understanding the development of readers' sensitivity to word endings as stress cues. Yet to date, the literature on word endings' role as stress cues in developing readers is fairly small (see also Burani et al., 2014; Colombo et al., 2014; Sulpizio & Colombo, 2013 for work in Italian) and warrants further investigation.

To address the important issue of children's developing sensitivity to the links between stress and spelling in English, I present cross-sectional data from elementary school-aged children in Chapters 4 and 5. Drawing from Arciuli and colleagues' (2010) empirical findings, I expected an effect of word endings as stress cues to emerge in the mid-elementary grades (starting in approximately Grade 3), and to increase through the late-elementary grades (Grade 6). That is, in Chapter 4's pseudoword spelling choice experiment, I expected students in the mid- and late-elementary grades (Grades 3 to 6) to choose spellings that are associated with the stress patterns of aurally presented pseudowords. However, I did not expect stress to guide spelling choices in younger, early-elementary students (Grades 1 and 2). In Chapter 5, I focused on children in the mid- to late-elementary grades to determine whether young readers use word endings as cues to stress in connected text. As with adults, I expected this effect to be modest, if present at all.

Together, the research presented in this dissertation will serve to clarify core theoretical and empirical questions about multisyllabic word reading in English. I present these experiments in Chapters 2 through 5, followed in Chapter 6 by a discussion of what these findings tell us about the nature and impact of word endings as cues to stress in multisyllabic written words.

Table 1.1.

A comparison of three models of disyllabic word reading, summarizing the characteristics of each model that inform the theoretical issues explored in this dissertation.

Theoretical issue	**DUAL-ROUTE:** [a] *Rastle & Coltheart (2000)*	**CONNECTIONIST:** *Ševa, Monaghan, & Arciuli (2009)*	**CDP++:** [a] *Perry, Ziegler, & Zorzi (2010)*
Roles of suffix and non-suffix endings as stress cues	- Conceptually, **affixes are central** to stress assignment. - Non-suffix endings are not included. - In practice, suffix and pseudosuffix endings are not differentiated (*beggar* vs. *mortar*).	- **Suffixes have no special status** as cues to stress. - Includes both suffix and non-suffix endings. No built-in differentiation between ending types. - Model identifies relevant cues from trained input.	- **Suffixes have no special status** as cues to stress. - Includes both suffix and non-suffix endings. No built-in differentiation between ending types. - Model identifies relevant cues from trained input.
Mechanism behind readers' learning of stress cues	- **Rule-based procedure** used to assign stress based on a stored list of affixes. - No indication of how readers might acquire rules to stress placement. Whether they do so at all is recognized as an assumption of the model.	- **Statistical learning mechanism**. - All regularities related to stress placement are learned from the training input. - Words presented during training are sampled according to their frequency in the language, mimicking readers' exposure.	- **Statistical learning mechanism**. - All regularities related to stress placement are learned from the training input. - Learning rates for words presented during training are scaled according to their frequency in the language, mimicking readers' exposure.
Approach to handling vowel quality in stress assignment	- Treats vowel quality and stress as **overlapping**. - Vowel reduction and stress assignment co-occur in several steps of the model's algorithm; the two cannot be separated.	- Model **does not include** information about vowel quality. Only maps orthography to stress. - No phoneme-level output; thus, no indication of how (or whether) the model links stress with vowel quality.	- Treats vowel quality and stress as **distinct**. - No connections between the model's phoneme nodes and stress nodes, at either the input level or the output level.
Reading performance captured by the model	- Reading **isolated words** and nonwords aloud. - Stress placement and segmental pronunciation accuracy are provided as output.	- Reading **isolated words** and nonwords aloud. - Stress placement accuracy is the only provided output.	- Reading **isolated words** and nonwords aloud. - Stress placement, segmental pronunciation, and reaction times are provided as output.

[a] Summaries describe these models' sublexical routes.

CHAPTER 2
SUFFIX AND NON-SUFFIX WORD ENDINGS AS CUES TO STRESS IN WRITTEN ENGLISH

2.1 ABSTRACT

Prior research shows that readers of English use both suffix and non-suffix word endings as cues to lexical stress. To date, however, these investigations have been separate, leaving open questions about the relative strength of the two ending types as stress cues. In the current study, we addressed this gap with two experiments in which English-speaking adults completed visual lexical decision tasks. Participants were quicker to respond to words whose endings accurately cue stress than words whose endings offer a misleading cue to stress. This effect was similar across endings that can and cannot be used as English suffixes (e.g., *–ar* and *–el*, respectively), suggesting that suffixes are not uniquely strong cues to stress in English words. Moreover, response latencies were similar regardless of whether endings served as meaningful suffixes (e.g., *beggar*) or as pseudo-suffixes (e.g., *mortar*). These findings differentiate between competing theoretical accounts of stress assignment, supporting a connectionist mechanism over a suffix-driven, rule-based one.

2.2 INTRODUCTION

In English, lexical stress placement is a challenge unique to the reading and recognition of multisyllabic words. However, the issue of how readers place stress in written words has traditionally been overlooked as a consequence of the field's focus on monosyllabic word reading (e.g., Coltheart, Curtis, Atkins, & Haller, 1993; Seidenberg & McClelland, 1989). Recently, there have been increasing empirical and theoretical efforts to understand multisyllabic word reading, including investigations of the sublexical cues that readers draw upon to guide stress placement (e.g., Arciuli, Monaghan, & Ševa, 2010; Perry, Ziegler, & Zorzi, 2010; Rastle & Coltheart, 2000; Ševa, Monaghan, & Arciuli, 2009). Written word endings (e.g., *–ar*; *–el*) act as one such sublexical cue—corpus analyses have shown that many endings' spellings are probabilistically associated with lexical stress patterns in English (Arciuli & Cupples, 2006). Building on this, separate lines of behavioural research have demonstrated that English-speaking adults draw on suffixes (e.g., Wade-Woolley & Heggie, 2015) and on non-suffix word endings (e.g., Arciuli & Cupples, 2006; Mundy & Carroll, 2013) as cues to stress placement. Yet crucially, no research to date has directly compared the extent to which these two ending types act as stress cues during visual word recognition. According to one prominent model of disyllabic word reading, suffixes play a uniquely strong role as cues to stress in English (Rastle & Coltheart, 2000). In direct contrast, connectionist models of disyllabic word reading argue that suffixes are not inherently stronger cues to stress than are non-suffix word endings (Arciuli et al., 2010; Ševa et al., 2009). In the current study, we offer empirical tests of these competing predictions.

2.2.1 SUFFIXES AS CUES TO STRESS IN ENGLISH

Strong linguistic and behavioural evidence points to a role for suffixes as cues to stress in English, as many derivational suffixes have predictable consequences on the stress patterns of words in which they appear. Some suffixes are stress-neutral, in that they do not affect the stress patterns of words to which they attach (e.g., *WONder–WONderful*; *HAppy–HAppiness*). Others are non-neutral, causing a shift in stress (e.g., *MAjor–maJORity*; *ATHlete–athLETic*). In either case, the suffix provides a clear source of information that readers can draw on to place stress in English words. This information is perhaps most evident in words with three or more syllables where the base itself is multisyllabic, though suffixes also inform stress placement in disyllabic words (e.g., *BEGGar*; Fudge, 1984).

In oral and written language tasks, English-speaking adults show clear sensitivity to the information that suffixes provide about stress placement. For instance, when asked to orally produce a derived word by adding a suffix to a stem (e.g., "Put *–ity* on the end of *active*"), adults' stress placement accuracy is at ceiling for words with both stress-neutral and stress-shifting suffixes (Jarmulowicz, 2006). Similarly, adults prefer derived pseudowords whose stress patterns align with the pattern suggested by their suffixes (e.g., *bisTINity* is preferred over *BIStinity*; preference for *FROsureful* over *froSUREful* depends on the stress pattern given for the stem *frosure*; Wade-Woolley & Heggie, 2015). Adults also apply their knowledge of suffix-based rules for stress assignment to reading tasks, producing stress in written pseudowords in accordance with their suffixes at above-chance levels (Wade-Woolley & Heggie, 2015). Further evidence suggests that, among both adults and children, the ability to use suffixes as cues to stress in English is positively related to general reading ability (e.g., Clin, Wade-Woolley, & Heggie, 2009; Jarmulowicz, Hay, Taran, & Ethington, 2008; Wade-Woolley & Heggie, 2015).

Drawing on the link between stress and suffixes in English, Rastle and Coltheart (2000) developed a model of disyllabic word reading that was designed to assign stress from a word's orthography. Their model adopted the rule-based architecture of prominent dual-route models of monosyllabic word reading (e.g., Coltheart et al., 1993), and they focused on creating a sublexical route to stress assignment that uses affixes as its foundation. Specifically, Rastle and Coltheart's (2000) model is based around a store of

54 English prefixes and 101 English suffixes. This affix store contains information about each affix's spelling and its effect on stress placement (based on Fudge, 1984). When faced with a written word, the model's rule-based algorithm recursively searches for prefixes and suffixes (along with checks for phonotactic and orthographic legality to ensure that removing the word's affixes leaves behind a plausibly-spelled English base word). It then uses the identified affixes' characteristics to assign stress and determine vowel quality. In words that do not contain an affix, the model defaults to assigning first-syllable stress, which is the most common stress pattern among disyllabic words in English (e.g., Kelly & Bock, 1988).

Two characteristics of Rastle and Coltheart's (2000) model are particularly relevant for our purposes. The first is that the model excludes non-suffix word endings almost entirely. Aside from a small number of non-suffix endings that are strongly associated with second-syllable stress (e.g., *–oo*; Fudge, 1984), affixes are the model's only link between a word's spelling and stress pattern. As a result, the model is not equipped to use non-suffix word endings as stress cues (e.g., *–el*). A strong interpretation of this design feature would suggest that suffixes are the only type of word endings to act as stress cues in English. Even a more moderate take on the model's architecture would suggest that suffixes should serve as stronger cues to stress than non-suffix word endings. To date, however, the premise that suffixes play a special role as cues to stress in English has not been tested empirically.

A second relevant characteristic of Rastle and Coltheart's (2000) model is that it does not differentiate between suffix and pseudo-suffix uses of word endings. Here, suffixes refer to word endings that act as meaningful units in the words in which they appear (e.g., the agentive suffix *–ar* in *beggar*), whereas pseudo-suffixes are endings that have the spelling of an English suffix but that do not act as meaningful units (e.g., the *–ar* in *mortar*). Strictly speaking, it is more precise to say that Rastle and Coltheart's model is built around "affix-like [letter] strings" (Rastle & Coltheart, 2000, p. 350) than around meaningful affixes. Thus, in the current study, we explore the claim that readers process all words with suffix-like endings—whether they are real suffixes or pseudo-suffixes—in a similar way.

2.2.2 NON-SUFFIX ENDINGS AS CUES TO STRESS IN ENGLISH

Suffixes are not the only type of word ending to provide information about stress placement. Indeed, corpus analyses have shown that many non-suffix word endings are probabilistically associated with stress patterns in English (Arciuli & Cupples, 2006; Arciuli et al., 2010; Kelly et al., 1998; Monaghan et al., 2016). An influential demonstration of this link between word endings and lexical stress comes from work by Arciuli and Cupples (2006). In their corpus analysis, endings were identified as the rime of a disyllabic word's second syllable (capturing the first phonemic vowel and all letters that follow), and endings' associations with stress were quantified as the proportion of words with an ending that have a given stress pattern. For example, roughly 89% of disyllabic words ending in *–ar* have first-syllable stress, as do 91% of words that end in *–el*. Notably, Arciuli and Cupples determined that only 14% of the endings included in their analysis can be used as English suffixes (e.g., *–ar*); the other 86% of endings cannot (e.g., *–el*).[2] This work helps to support the claim that non-suffix word endings can act as

[2] Note that 14% likely underestimates the proportion of English suffixes captured in this corpus analysis, as Arciuli and Cupples' (2006) definition of word endings overlooks suffixes that begin with consonants. For instance, the ending *–ul* would not be identified

effective, though probabilistic, cues to stress in English words.

Behavioural evidence suggests that English-speaking adults use the associations between non-suffix endings and lexical stress when processing written words. They respond more quickly to words whose endings provide accurate cues to stress than words whose endings provide misleading cues to stress—an effect that has been seen in both naming tasks (Kelly et al., 1998) and lexical decision (Kelly et al., 1998; Mundy & Carroll, 2013). These findings suggest that adults' reading and recognition of familiar words is facilitated when non-suffix endings cue a stress pattern that is consistent with a word's actual stress pattern (e.g., the ending *–el* acts as a cue to first-syllable stress in *NOVel*). In contrast, written word recognition is hindered when non-suffix endings are inconsistent with a word's stress pattern (e.g., *–el* in *hoTEL*). Similar behavioural findings show that adults use non-suffix endings to place stress in unfamiliar words. For instance, when presented with disyllabic written pseudowords and asked to indicate which syllable should receive emphasis, participants tend to assign the stress pattern associated with each pseudoword's ending (e.g., *RANcel*; *feDUCT*; Arciuli & Cupples, 2006). Taken together, this research demonstrates that adults are sensitive to the information about stress placement that non-suffix endings can provide.

Building on this corpus and behavioural evidence, connectionist models of disyllabic word reading have been developed as theoretical alternatives to Rastle and Coltheart's (2000) rule-based model (Arciuli et al., 2010; Perry et al., 2010; Ševa et al.,

---

as an English suffix (per Fudge, 1984), though many of the words in which it appears end with the suffix *–ful*. However, a later corpus analysis of children's texts supports the claim that non-suffix word endings are associated with stress, even after excluding all morphological influence (Arciuli et al., 2010). In that study, results were nearly identical when the corpus analysis was based on all disyllabic words and when it was repeated based on monomorphemic words alone.

2009). An early model created by Ševa and colleagues (2009) is a notable example. Their model aims to characterize the process by which readers learn of, and subsequently use, the links between stress and spelling in English. Notably, it does so solely through exposure to written words. There is no store of affixes to provide linguistically-driven rules for stress assignment (cf. Rastle & Coltheart, 2000); in fact, the model does not specify any cues to stress *a priori*. Instead, the model's knowledge of written stress cues emerges through a training process that mirrors readers' exposure to text (Ševa et al., 2009).

Through this training process, Ševa and colleagues' (2009) model is able to learn of word endings' associations with stress—regardless of the endings' morphological status. As the authors put it, "[morphological] cues are part of the orthographic information, but, in the connectionist tradition, they do not have special status beyond other, equally reliable cues that may be present … to determine stress position" (p. 241). In other words, under a connectionist approach to stress assignment, suffix-like endings and non-suffix endings should have similar effects on visual word recognition. This prediction seems plausible, given the behavioural evidence for readers' sensitivity to non-suffix endings as stress cues (e.g., Arciuli & Cupples, 2006; Kelly et al., 1998; Mundy & Carroll, 2013). However, it stands in stark contrast with the predictions of Rastle and Coltheart's (2000) model, and without a direct comparison of suffix-like and non-suffix endings, we cannot draw firm conclusions about their relative strengths as stress cues. In the current study, we provide this direct comparison.

2.2.3 THE CURRENT STUDY

In sum, there is compelling behavioural evidence that English-speaking adults recognize that both suffix and non-suffix word endings provide information about stress placement in written words (e.g., Kelly et al., 1998; Wade-Woolley & Heggie, 2015). However, the relative strength of the two ending types as cues to stress is an open question. This empirical gap represents a key point of contrast between existing theories of disyllabic word reading. Broadly, the rule-based account put forward by Rastle and Coltheart (2000) predicts a stronger role for suffix-like word endings (which have the spelling of an English suffix, regardless of whether the ending serves as a meaningful unit) than for non-suffix word endings (whose spellings cannot be used as English suffixes). In contrast, connectionist models predict that the two ending types will have similar effects on word reading and recognition; under this account, suffixes play no special role as cues to stress (Arciuli et al., 2010; Ševa et al., 2009). In the current study, we provide empirical tests of these competing predictions.

An ideal test of these predictions would involve a full factorial manipulation of words with suffixes, pseudo-suffixes, and non-suffix endings that are either consistent or inconsistent with the stress patterns of the words in which they appear. However, constraints of the language introduce logistical challenges to that ideal. Specifically, there are exceptionally few words with true suffixes that provide misleading cues to stress (e.g., all disyllabic English words in which *–ar* appears as a suffix have first-syllable stress); this makes an inconsistent suffix condition impossible. As a solution to this challenge, we conducted two experiments: one comparing words with pseudo-suffix and non-suffix endings, and a second comparing words with suffix and pseudo-suffix endings. In both experiments, English-speaking adults completed a visual lexical decision task. Although the existing models of disyllabic word reading focus on stress placement when reading aloud, their predictions can be applied conceptually to the process by which readers assign stress during visual word recognition (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). For our purposes, lexical decision offers a key advantage over tasks that involve reading words aloud; evidence suggests that disyllabic words' stress patterns affect naming latencies, but stress does not reliably affect lexical decision latencies (e.g., Yap & Balota, 2009).

In Experiment 1, we manipulated words' ending type (pseudo-suffix; non-suffix) and stress consistency (consistent; inconsistent) factorially. Both theoretical models would predict a main effect of endings' stress consistency, with a behavioural advantage for words whose endings accurately cue their stress pattern (consistent condition; e.g., *MORTar*; *NOVel*) over words whose endings do not (inconsistent condition; e.g., *guiTAR*; *hoTEL*). Crucially, though, the models make differing predictions about the interaction between ending type and stress consistency. Under Rastle and Coltheart's (2000) rule-based account, we would expect to see a stronger effect of endings' consistency with stress for words with pseudo-suffixes than words with non-suffix endings. In contrast, under Ševa and colleagues' (2009) connectionist account, we would not expect an interaction between ending type and ending stress consistency; the effect of stress consistency should be similar across pseudo-suffixed and non-suffixed words. Despite the lack of items with real suffixes in Experiment 1, using pseudo-suffixed items allows us to test the underlying predictions of the two competing models—neither differentiates between real suffix and pseudo-suffix uses of word endings (Rastle & Coltheart, 2000; Ševa et al., 2009).

In Experiment 2, we compared adults' lexical decisions toward words with true English suffixes (e.g., *BEGGar*) and words in which the same endings are pseudo-suffixes (e.g., *MORTar*). In both cases, these endings are consistent with their words' stress patterns. Although this comparison cannot test the critical interaction laid out in Experiment 1, it offers a useful test of the models' decision to treat all suffix-like word endings as similar. Finding a difference in readers' responses to the two ending types (in either direction) would be particularly problematic for Rastle and Coltheart's model, as it explicitly treats all suffix-like word endings as equivalent, regardless of their morphological status. However, finding no difference between suffixed and pseudo-suffixed words would help to validate this characteristic of their model. In addition to the ending type manipulation, Experiment 2 includes an inconsistent pseudo-suffix condition (e.g., *guiTAR*); this allows for a second test of the claim that word endings' consistency with stress affects readers' lexical decisions. As in Experiment 1, we expect to see faster responses toward words with consistent pseudo-suffix endings than words with inconsistent pseudo-suffix endings (e.g., Kelly et al., 1998; Mundy & Carroll, 2013).

2.3 EXPERIMENT 1

2.3.1 METHOD

Participants. Participants were 60 university students aged 17 and older (39 women; mean age = 20.76 years; *SD* = 1.71 years). All participants spoke English as a first language and had normal or corrected-to-normal vision. Two additional participants were tested, but were excluded from analyses due to low accuracy scores.[3] As

[3] Across items, the excluded participants had mean accuracy rates of 10.9% and 31.2%, respectively. We suspect that these accuracy rates are due to experimenter error in matching the response keys to the participants' handedness.

compensation, participants earned either $5 or 0.5 percentage points towards a university course.

Materials. Participants completed a visual lexical decision task consisting of 64 real word items and a corresponding 64 pseudoword items. All real word items used in this experiment were disyllabic, monomorphemic English words (per the CELEX English database; Baayen et al., 1995). They were divided into one of four conditions resulting from the factorial manipulation of ending type (pseudo-suffix; non-suffix) and ending stress consistency (consistent; inconsistent). These items are listed in Appendix A.

Our ending type manipulation was based on work by Fudge (1984), which provides a comprehensive list of English affixes and their stress-related properties. This same list was used as the basis for Rastle and Coltheart's (2000) rule-based stress assignment algorithm, making it an ideal resource with which to test predictions of their model. Endings in our pseudo-suffix condition were among those included in Fudge's (1984) list of English suffixes. Although these pseudo-suffix endings can be used as meaningful English suffixes, they are not used as such in any of Experiment 1's items (e.g., *–ar* is not a suffix in *mortar*). By contrast, endings in our non-suffix condition were not included in Fudge's (1984) list of suffixes, nor did they appear as suffixes among the disyllabic words in the CELEX English database (Baayen et al., 1995).[4] Conditions were balanced on the strength of their endings' associations with stress, as measured by the proportion of first-syllable stressed words among all disyllabic words with a given ending

[4] Wherever possible, we also ensured that our non-suffix endings did not occur as part of a longer English suffix, though there were two exceptions to this criterion (*–ard* and *–one*, which appear in the suffixes *–ward* and *–phone*, respectively; Fudge, 1984). To avoid potential issues with these endings, our lexical decision task did not contain items ending with *–ward* or *–phone*.

(pseudo-suffix endings: *M* = .86, *SD* = .09; non-suffix endings: *M* = .83, *SD* = .06), $t(18) = 0.97$, $p = .35$.

Next, we selected real-word items whose stress patterns were either consistent or inconsistent with the stress pattern suggested by their endings. Items were selected in pairs that shared an ending; this ensured that the same pseudo-suffix and non-suffix endings were used in their respective consistent (e.g., *PILLar*; *CHAPel*) and inconsistent (e.g., *guiTAR*; *pasTEL*) conditions. We included 16 items per condition for a total of 64 real word items. Items with alternate spellings or stress patterns were excluded from consideration. Endings appeared in 1–2 words per condition.

It is worth noting a potential confound in our design—words in the consistent condition all had first-syllable stress whereas those in the inconsistent condition all had second-syllable stress. This confound results from the fact that few of the English suffixes used in disyllabic words are stress-taking (Fudge, 1984); as such, they appear primarily in words with first-syllable stress. And although some English suffixes are indeed associated with second-syllable stress, very few of those endings lend themselves to our stress consistency manipulation. Consequently, all endings used in our task are primarily associated with first-syllable stress. Despite this fact, we argue that our study's design is justified by prior research showing that adults' lexical decision latencies toward disyllabic words are not reliably affected by differences in stress (e.g., Mundy & Carroll, 2013; Yap & Balota, 2009). Moreover, even if this prior work has missed an underlying effect of stress patterns on lexical decision latencies, we would expect it to apply to both our pseudo-suffix and non-suffix ending conditions. As such, any potential confound should not affect the critical interaction between ending type and ending stress consistency that Experiment 1 was designed to test.

Our four experimental conditions were balanced on word frequency (p = .73; occurrences per million, Baayen et al., 1995), word length (p = .33), orthographic Levenshtein distance (p = .77; Balota et al., 2007), and mean bigram frequency (p = .56; Balota et al., 2007). Descriptive statistics for these lexical characteristics are summarized by condition in Table 2.1. We also matched conditions on the number of words that began with a pseudo-prefix spelling according to Fudge's (1984) list of prefixes and the rules implemented by Rastle and Coltheart's (2000) algorithm.[5] This matching was done because their model uses prefix-like spellings as a central cue to stress placement alongside suffix-like cues. None of our items began with a real English prefix (per Baayen et al., 1995).

---

[5] Specifically, Rastle and Coltheart's (2000) model considers a word prefixed if (1) it begins with a spelling sequence that appears in Fudge's (1984) list of English prefixes, and (2) the prefix-like spelling is followed by an orthographically legal word-initial bigram. For example, *re–* would be classified as a prefix in *regain* and *regent*, but not in *rental*.

Table 2.1.

Descriptive statistics by condition, summarizing the lexical characteristics on which Experiment 1's conditions were balanced.

	Lexical characteristics: *M* (*SD*)				# of pseudo-prefixed items
Condition	Word frequency	Word length	OLD	Mean bigram frequency	
Pseudo-suffix ending					
Consistent	44.1 (78.8)	6.3 (1.1)	2.2 (0.4)	1974.4 (671.7)	6
Inconsistent	39.6 (60.0)	6.0 (0.7)	2.2 (0.4)	1729.2 (529.9)	6
Non-suffix ending					
Consistent	19.0 (48.8)	5.7 (1.0)	2.2 (0.4)	1702.0 (606.5)	6
Inconsistent	36.4 (72.4)	6.1 (1.0)	2.3 (0.5)	1785.0 (516.4)	6

Note. OLD = Orthographic Levenshtein Distance.

As foils for our lexical decision task, we created 64 pseudoword items by changing one or two letters in each real word target item (e.g., Ford, Davis, & Marslen-Wilson, 2010). Pseudowords contained the same ending spellings as the real word items, and we created the pseudowords with the intent that participants would read them as disyllabic (cf. Kelly et al., 1998). All pseudowords followed the orthographic and phonotactic constraints of English and maintained the same number of letters and phonemes as their real-word counterparts. As such, our pseudoword items were highly similar to our real word targets, resulting in a relatively high-demand lexical decision task (e.g., Stone & Van Orden, 1993; Yang & Zevin, 2014).

Procedure. Participants were tested individually on a PC laptop running DirectRT software (Jarvis, 2008a). A fixation cross was presented for 1000ms at the start of each trial, followed by a word or pseudoword item (Times New Roman font, size 40).

Participants were asked to decide if the presented item was a real English word; they then responded by pressing a green “Y” button if the item was a real word or a red “N” button if it was not. The position of the response keys was adjusted so that participants indicated real words with their dominant index finger and pseudowords with their nondominant index finger. On each trial, participants were given up to 5000ms to respond, at which point the target disappeared. Participants completed eight practice trials, during which the experimenter provided corrections for miscategorised words. They then completed the 128 experimental trials, which were presented in a randomized order without feedback. The task took approximately 10–15 minutes to complete.

2.3.2 RESULTS

Data cleaning. Prior to analysis, we inspected accuracy across items, participants, and conditions. None of our items had problematically low accuracy rates (i.e., none fell significantly below a mean accuracy rate of 50%); we therefore retained all items for analysis (e.g., Jiang, 2012). As noted in section 2.3.1, two participants were excluded due to their low accuracy scores; however, all remaining participants had mean accuracy rates well above chance level. In each of our four experimental conditions, accuracy scores were at ceiling (see Table 2.2). As such, our analyses focus on participants’ reaction time to correctly identified words. We examined the reaction time data for extreme values and outliers. There were no extreme values requiring exclusion (i.e., no values were longer than 5000ms or shorter than 250ms). However, we removed 114 data points as outliers (2.96% of all data), defined here as responses that fell more than 2.5 standard deviations from each participant’s mean (cutoff selected based on Van Selst & Jolicoeur, 1994; non-recursive moving criterion method). Table 2.2 shows participants’ mean reaction times by condition.

Table 2.2.

Mean reaction times to correctly identified words and mean accuracy rates (with standard deviations) for Experiment 1.

	Reaction time (ms)		Accuracy (% correct)	
Condition	Mean	*SD*	Mean	*SD*
Pseudo-suffix ending				
Consistent	626.70	104.90	87.5%	33.1%
Inconsistent	647.17	104.70	89.1%	31.2%
Non-suffix ending				
Consistent	631.78	80.36	86.3%	34.5%
Inconsistent	642.11	108.92	83.5%	37.1%

Our analyses used log-transformed reaction time to correctly identified words as their outcome measure. Log transformation successfully addressed the positive skew that was present in the residuals when raw reaction time data were analyzed (Tabachnick & Fidell, 2007). However, note that the pattern of results that we report below holds true when analyzing both raw and log-transformed reaction times.

Analysis. To identify the ending characteristics that affect readers' lexical decision latencies, we conducted a 2 x 2 repeated-measures ANOVA with endings' stress consistency (consistent; inconsistent) and ending type (pseudo-suffix; non-suffix) as within-subjects factors. As expected, the main effect of endings' stress consistency was significant, $F(1, 59) = 4.48$, $p = .039$, $\eta_p^2 = 0.07$; participants responded more quickly to words whose endings were consistent with their stress pattern (log-transformed reaction time: $M = 2.795$, $SD = 0.057$) than words whose endings were inconsistent with their stress pattern (M = 2.803, SD = 0.062). However, the main effect of ending type was nonsignificant, $F(1, 59) = 0.05$, $p = .823$, $\eta_p^2 = 0.001$ (pseudo-suffix: M = 2.799, SD = 0.066; non-suffix: M = 2.798, SD = 0.059), as was the interaction between ending type and stress consistency, $F(1, 59) = 2.08$, $p = .154$, $\eta_p^2 = 0.03$. Figure 2.1 in Appendix B displays these key effects, showing paired differences in participants' mean response times toward words whose endings were consistent versus inconsistent with their stress patterns. Negative mean scores illustrate the main effect of endings' stress consistency. These paired difference scores are presented as a function of ending type, illustrating the null interaction.

Following on these findings, we conducted a parallel item analysis with endings' stress consistency (consistent; inconsistent) and ending type (pseudo-suffix; non-suffix) as between-subjects factors. Neither main effect was significant, nor was the interaction between ending type and stress consistency; Fs(1, 60) < 0.76, ps > .39, η_p^2s < .013. However, the pattern of means aligned with the patterns reported in our analysis by participants. That is, responses were slightly faster toward items whose endings were consistent (log-transformed reaction times: M = 2.812, SD = 0.042) versus inconsistent (M = 2.822, SD = 0.043) with their stress pattern. Response latencies were similar for words with pseudo-suffix (M = 2.817, SD = 0.048) and non-suffix endings (M = 2.816, SD = 0.044), and there was no apparent trend toward an interaction (consistent pseudo-suffix: M = 2.811; SD = 0.036; inconsistent pseudo-suffix: M = 2.824; SD = 0.049; consistent non-suffix: M = 2.813; SD = 0.049; inconsistent non-suffix: M = 2.819; SD = 0.039).

Taken together, Experiment 1 allows us to draw two key conclusions. First, its findings confirm that English-speaking adults are quicker to respond to words whose endings are consistent with their stress pattern than to words whose endings are inconsistent with their stress pattern (see also Kelly et al., 1998; Mundy & Carroll, 2013). Second, we show that the effect of endings' stress consistency is similar across pseudo-suffix and non-suffix endings. However, Experiment 1 did not include words with meaningful suffixes, and it therefore cannot speak to whether readers process words with suffix and pseudo-suffix endings similarly. We address this question in Experiment 2.

2.4 EXPERIMENT 2

2.4.1 METHOD

Participants. Participants were 60 university students aged 17 and older (45 women; mean age = 19.49 years; *SD* = 1.14 years). All spoke English as a first language and had normal or corrected-to-normal vision; none took part in Experiment 1. As compensation, participants earned either $5 or 0.5 percentage points toward a university course.

Materials. Real word item selection and pseudoword item generation followed the criteria described in Experiment 1. The key difference between experiments was that Experiment 2 focused solely on suffix-like ending spellings; it did not include non-suffix conditions. Instead, our items fell into one of three conditions: (1) suffixed words whose endings are consistent with their stress pattern, (2) pseudo-suffixed words whose endings are consistent with their stress pattern, and (3) pseudo-suffixed words whose endings are inconsistent with their stress pattern. We selected items in sets that shared an ending; for example, *BEGGar* (consistent suffix), *MORTar* (consistent pseudo-suffix), and *guiTAR* (inconsistent pseudo-suffix). All endings appeared in Fudge's (1984) list of English suffixes, and a word's status as either suffixed or pseudo-suffixed was confirmed using morphological status designations in the CELEX English database (Baayen et al., 1995). The items used in Experiment 2 are listed in Appendix C.

In this experiment, we were not able to include an inconsistent suffix condition due to a lack of viable items. For instance, none of the disyllabic English words in which *–ar* acts as a suffix are pronounced with second-syllable stress. This constraint in the language prevented us from re-creating the factorial ending type x stress consistency design used in Experiment 1. However, Experiment 2's design allowed us to make two separate comparisons. The first comparison involves our two pseudo-suffix conditions (consistent and inconsistent), aiming to replicate the stress consistency effect that emerged in Experiment 1. The second involves our consistent suffix and consistent pseudo-suffix conditions. This comparison allows us to either validate or refute Rastle and Coltheart's (2000) decision to treat all suffix-like endings as the same in their model.

Experiment 2's lexical decision task consisted of 96 real word items (64 target words plus 32 filler words, which are described below) and 96 pseudoword items (generated according to the criteria listed in Experiment 1). Appendix C shows the division of the target words into our three conditions. The inconsistent pseudo-suffix condition contains 16 items (each of which is matched to an item in the consistent pseudo-suffix condition against which it is being compared). However, the consistent pseudo-suffix and consistent suffix conditions each have 24 items in total. Sixteen of these items share endings with the inconsistent pseudo-suffix items. The additional 8 items have endings that always (or nearly always) appear in words with first-syllable stress (e.g., *–ic*; *–ish*). We added these items because their near-deterministic associations with stress make them ideally suited to the rule-based stress assignment procedure used in Rastle and Coltheart's (2000) model. Given that our comparison of suffixed and pseudo-suffixed items is fundamentally a test of Rastle and Coltheart's decision to treat all suffix-like endings as similar, we felt that it was important to include items whose endings act as particularly strong cues to stress.

Our three experimental conditions were balanced on word frequency (*p*s > .58), word length (*p*s > .38), orthographic Levenshtein distance (*p*s > .16), and mean bigram frequency (*p*s > .80). Descriptive statistics for these lexical characteristics are summarized in Table 2.2. As in Experiment 1, conditions were matched on the number of items that began with a pseudo-prefix spelling (Fudge, 1984), and no items began with a genuine prefix. Because a majority of our target items have first-syllable stress, we included 32 filler words with second-syllable stress. These filler items were not included in our analyses.

Procedure. As described in Experiment 1. In total, the lexical decision task took approximately 15–20 minutes to complete.

Table 2.3.

Descriptive statistics by condition, summarizing the lexical characteristics on which Experiment 2's conditions were balanced.

Condition	Lexical characteristics, *M* (*SD*) Word frequency	Word length	OLD	Mean bigram frequency	# of pseudo-prefixed items
Stress consistency comparison					
Consistent [a]	44.0 (78.8)	6.1 (1.0)	2.2 (0.4)	1805.2 (595.1)	3
Inconsistent [a]	31.3 (57.3)	6.0 (0.8)	2.2 (0.4)	1850.2 (637.1)	3
Ending type comparison					
Suffix [b]	46.3 (73.1)	6.5 (1.0)	2.5 (0.5)	1855.1 (599.6)	3
Pseudo-suffix [b]	34.9 (66.0)	6.3 (1.0)	2.3 (0.5)	1897.7 (574.6)	3

Notes. OLD = Orthographic Levenshtein Distance. [a] Both are pseudo-suffix conditions (16 items). [b] Both are consistent conditions (24 items).

2.4.2 RESULTS

Data cleaning. Data were prepared for analysis using the procedures described in Experiment 1. None of the items or participants had problematically low accuracy rates (i.e., below 50% chance); all were retained in our analysis. Across conditions, accuracy scores were at ceiling (see Table 2.4). Our analyses therefore focus on reaction time to correctly identified words. We excluded 5 reaction times that were longer than 5000ms (0.001% of all data) and excluded 127 as outliers that fell more than 2.5 standard deviations above each participant's mean (3.31% of all data). There were no overly short reaction times requiring exclusion. Table 2.4 shows participants' mean reaction times by condition. As in Experiment 1, we transformed reaction time scores logarithmically to address positive skew (Tabachnick & Fidell, 2007). We excluded one participant as a univariate outlier, as their reaction times were extreme in each of the experimental conditions, regardless of the transformations attempted on the data (with z-scores ranging from 3.41 to 4.55 in the log-transformed RT data). Note that the pattern of results reported below is not changed by the inclusion or exclusion of this participant.

Table 2.4.

Mean reaction times to correctly identified words and mean accuracy rates (with standard deviations) for Experiment 2.

	Reaction time (ms)		Accuracy (% correct)	
Condition	Mean	*SD*	Mean	*SD*
Stress consistency comparison				
Consistent [a]	656.26	129.52	93.6%	24.4%
Inconsistent [a]	698.56	150.46	89.2%	31.3%
Ending type comparison				
Suffix [b]	675.81	137.40	92.6%	26.2%
Pseudo-suffix [b]	662.12	130.33	92.9%	25.8%

Notes. [a] Both are pseudo-suffix conditions (16 items). [b] Both are consistent conditions (24 items).

Analyses. Our first analysis sought to replicate the stress consistency effect that we found in Experiment 1 (see also Kelly et al., 1998; Mundy & Carroll, 2013). To that end, we conducted a paired-samples t-test comparing response latencies toward words in our two pseudo-suffix conditions. This analysis confirmed that readers responded more quickly to words whose endings are consistent with their stress pattern than to words whose endings are inconsistent with their stress pattern, $t(58) = 4.10$, $p < .001$, $d = 0.54$. The left-hand panel of Figure 2.2 (Appendix B) shows paired difference scores illustrating this effect of endings' stress consistency. A by-items analysis showed the same numerical pattern, as responses were somewhat faster toward words whose endings are consistent (*M* = 2.821, *SD* = 0.044) than inconsistent (*M* = 2.848, *SD* = 0.054) with their stress pattern. However, this pattern was nonsignificant by items, $t(30) = 1.53$, $p = .137$, $d = 0.55$.

Turning to ending type, our second analysis compared response times toward words whose endings are suffixes (e.g., *BEGGar*) and words in which those same endings are pseudo-suffixes (e.g., *MORTar*; recall that in this comparison, both conditions' endings are consistent with their stress patterns). A paired-samples *t*-test showed no difference in lexical decision latencies between the suffix and pseudo-suffix conditions, $t(58) = 1.23$, $p = .22$, $d = 0.16$. The right-hand panel of Figure 2.2 (Appendix B) illustrates this null effect of ending type. Likewise, we found no difference between the suffix and pseudo-suffix conditions in an analysis by items, $t(46) = 0.28$, $p = .778$, $d = 0.08$.

2.5 GENERAL DISCUSSION

Despite mounting evidence that adults use suffixes (e.g., Wade-Woolley & Heggie, 2015) and non-suffix endings (e.g., Kelly et al., 1998) as cues to stress in written English words, questions remain about the relative strength of these two types of ending cue. In the current study, we addressed this gap by comparing the speed of adults' lexical decisions toward words with suffixes, pseudo-suffix endings, and non-suffix endings. Our findings speak to an issue that is at the heart of competing models of disyllabic word reading—namely, whether suffix-like endings play a special role as cues to stress when compared with non-suffix endings (Rastle & Coltheart, 2000; Ševa et al., 2009). In Experiment 1, we compared readers' lexical decisions toward words with pseudo-suffix endings (e.g., *MORTar*; *guiTAR*) and those with non-suffix endings (e.g., *NOVel*;

hoTEL). In Experiment 2, we contrasted words with genuine suffixes (e.g., *BEGGar*) and words with pseudo-suffix endings (e.g., *MORTar*; *guiTAR*).

In both experiments, word endings' consistency with stress affected participants' lexical decision latencies. Specifically, adults were quicker to respond to words in the consistent conditions (where endings accurately cued words' stress patterns), and were slower to respond to words in the inconsistent conditions (where endings gave misleading cues to stress). These results add to two prior studies showing that English-speaking adults draw on written word endings as cues to stress when making lexical decisions (Kelly et al., 1998; Mundy & Carroll, 2013). They also align with related evidence that word endings affect adults' stress placement when reading pseudowords and familiar English words (e.g., Arciuli & Cupples, 2006; Kelly et al., 1998). Although the conclusion that word endings serve as cues to stress during visual word recognition is not unique to this study, our stress consistency analyses provide valuable corroborating evidence that readers are sensitive to the links between ending spellings and lexical stress in English—a claim that is central to theoretical models that assign stress from a word's orthography (e.g., Rastle & Coltheart, 2000; Ševa et al., 2009).

Interestingly, neither of this study's experiments found an effect of ending type on adults' lexical decision latencies. In Experiment 1, response times were similar toward words with pseudo-suffix and non-suffix endings; likewise, in Experiment 2, response times were similar toward words with suffix and pseudo-suffix endings. The latter of these findings warrants particular comment, as it contrasts with prior research showing that morphemic structure can affect visual word recognition among skilled and developing readers (e.g., Carlisle & Stone, 2005; Carlisle, Stone, & Katz, 2001; Casalis, Quémart, & Duncan, 2015; Deacon, Parrila & Kirby, 2006). When interpreting this contrast, it is important to note that the direction and magnitude of morphemic effects on word recognition can vary based on a range of factors that were not pertinent to our manipulation in the current study. For example, typical adult readers are quicker to recognize derived words than pseudo-derived words when no orthographic change occurs from a word's base to its derived form (e.g., *farmer* vs. *offer*), but this pattern is reversed in the presence of an orthographic change (e.g., *beginner* vs. *hammer*; Deacon et al., 2006). Similarly, adults' visual word recognition is slowed when derived words involve a phonological change from the base to derived form (e.g., *natural*) relative to derived words with no phonological change (e.g., *cultural*; Carlisle et al., 2001). Among developing readers of English, the presence of a base word (as in, e.g., *pressure*) slows visual word recognition when compared with words that lack a genuine base (as in e.g., *measure*; Casalis et al., 2015).

Although these distinctions are important to our overall understanding of visual word recognition, they are not relevant to the model algorithms that inspired the current study's manipulations (Rastle & Coltheart, 2000). As a result, our suffix condition included a mixture of items with orthographic and/or phonological changes (e.g., *beggar*; *tension*; *rhythmic*) and items with no such change (e.g., *movement*; *deeply*; *farmer*). Likewise, our pseudo-suffix conditions included a mixture of items with and without real English bases (e.g., *pill* in *pillar* and **oint* in *ointment*, respectively). The contrasting effects of these various item characteristics may have cancelled out any overall difference between our suffix and pseudo-suffix conditions in Experiment 2. Given these features of our items, we certainly do not present our ending type findings as a systematic examination of morphology's role in visual word recognition. Instead, we focus our interpretation around this study's main goal—providing empirical tests of the claims made by models of disyllabic word reading (Rastle & Coltheart, 2000; Ševa et al., 2009). With that in mind, Experiment 2 shows that there is no difference in the speed of adults' lexical decisions toward words with real suffix and pseudo-suffix endings when items are selected according to the criteria laid out in Rastle and Coltheart (2000). This finding validates their decision to treat all words with suffix-like endings as similar in their model.

The final salient finding from this study is Experiment 1's nonsignificant interaction between ending type and stress consistency. That is, the effect of word endings' consistency with stress on lexical decision latencies was similar for words with pseudo-suffix endings (e.g., *MORTar*; *guiTAR*) and those with non-suffix endings (e.g., *NOVel*; *hoTEL*). We argue that this finding represents the study's key theoretical contribution, and we elaborate on its implications for models of disyllabic word reading in more detail below.

2.5.1 THEORETICAL CONTRIBUTIONS

Our study's nonsignificant interaction between ending type and stress consistency presents a challenge for Rastle and Coltheart's (2000) affix-based model of disyllabic word reading. Specifically, in Experiment 1, word endings' consistency with stress had similar effects on lexical decision latencies across words with pseudo-suffix and non-suffix endings. This finding certainly does not refute the clear linguistic and behavioural evidence that suffixes provide useful cues to stress in English words (e.g., Fudge, 1984; Jarmulowicz, 2006; Wade-Woolley & Heggie, 2015). Instead, it demonstrates that suffix-like word endings do not play a uniquely strong role as cues to stress. As such, we argue that an effective model of disyllabic word reading should be able to use both ending types as cues to stress in written words. With this in mind, models that focus only on suffix-like word endings, as Rastle and Coltheart's model does, are incomplete. Conceivably, the stress assignment rules that drive Rastle and Coltheart's model could be modified to include non-suffix endings, though doing so would require treating the probabilistic associations between endings and stress (Arciuli & Cupples, 2006) as if they were deterministic cues. Further work is therefore needed to establish whether an adapted rule-based model can account for our empirical findings.

On the other hand, our nonsignificant interaction is fully compatible with Ševa and colleagues' (2009) connectionist model of disyllabic word reading. Their approach contends that all written word endings—regardless of their morphological status—can serve as probabilistic stress cues, with the key factor in a cue's utility being the strength of its association with stress. In support of this claim, our data affirm prior evidence that readers draw on non-suffix endings as cues to stress (e.g., Arciuli & Cupples, 2006), and we add to that work with a direct demonstration that non-suffix ending cues function much like endings that can be used as English suffixes. A theoretically parsimonious interpretation of this finding is that readers' stress assignment is driven by probabilistic cues that apply across word endings, rather than by formal stress-related properties of English suffixes[6] (Arciuli et al., 2010; Perry et al., 2010; Ševa et al., 2009).

[6] This interpretation aligns with the conclusions of recent work by Grimani and Protopapas (2016), who had Greek-speaking adults read pseudowords containing suffix-like endings. Linguistically, all suffixes used in their study had formal, predictable effects on stress placement (e.g., the suffix *–άδα*, which pulls stress to the syllable in which it appears; *πορτοκαλάδα*). However, while some endings were deterministically associated

2.5.2 CONCLUSIONS

At the core of this study is one theoretically-motivated question: do suffixes (or more precisely, suffix-like word endings, which have the spelling of an English suffix) play a special role as cues to stress in English words? Our results suggest that they do not. In Experiment 1, adults' lexical decision latencies were influenced by word endings' consistency with stress, but this effect was similar for words with pseudo-suffix and non-suffix endings. Further, in Experiment 2, responses were similar toward words with real suffix and pseudo-suffix endings. Together, these findings confirm that non-suffix endings act as valuable cues to stress—despite their probabilistic links with stress in English (e.g., Arciuli & Cupples, 2006), non-suffix endings' effect on visual word recognition is as strong as that of suffix-like endings. As with all empirical research, our conclusions apply within the bounds of our study's design, which focused solely on disyllabic words and faced constraints of the language relating to the availability of suffixed words. Even so, our findings effectively differentiate between competing theories of how stress is placed during disyllabic word reading, supporting a connectionist approach (Ševa et al., 2009) over a suffix-driven, rule-based one (Rastle & Coltheart, 2000). Taken together, the findings inform our understanding of the nature of written word endings as cues to lexical stress in English.

---

with stress (i.e., no words in Greek act as exceptions to the link between suffix and stress), others were less strongly associated with stress (due to words with pseudo-suffix uses of the ending that have a differing stress pattern). When pseudowords were presented without a diacritic, participants placed stress in accordance with the ending. However, ending-consistent stress placement was below ceiling even when endings were deterministically associated with stress (M = 76%), and there were significantly fewer ending-consistent responses when endings had only probabilistic associations with stress (M = 64%).

2.6 APPENDIX A: EXPERIMENT 1 ITEMS

A list of the real word items used in Experiment 1, presented by condition.

PSEUDO-SUFFIX		NON-SUFFIX	
Consistent	**Inconsistent**	**Consistent**	**Inconsistent**
pillar	*cigar*	*chapel*	*hotel*
mortar	*guitar*	*camel*	*pastel*
metal	*cabal*	*placard*	*bombard*
loyal	*banal*	*orchard*	*petard*
courage	*garage*	*shallow*	*endow*
adage	*montage*	*window*	*allow*
office	*suffice*	*ozone*	*condone*
lattice	*police*	*hormone*	*trombone*
congress	*success*	*cortex*	*perplex*
mattress	*duress*	*carol*	*patrol*
ointment	*augment*	*errand*	*command*
pigment	*lament*	*tofu*	*haiku*
future	*mature*	*lilac*	*shellac*
stature	*secure*	*oboe*	*canoe*
rally	*supply*	*format*	*cravat*
belly	*imply*	*debit*	*commit*

2.7 APPENDIX B: SUPPLEMENTAL FIGURES OF KEY STUDY FINDINGS

The figures in this appendix illustrate the effects found in Experiment 1 (Figure 2.1) and Experiment 2 (Figure 2.2), respectively. We include them to complement the descriptive statistics reported in text (Tables 2.2 and 2.4), which provide overall mean (and *SD*) reaction times by condition. In contrast to those condition means, the figures presented here plot paired difference scores to show the within-subject differences in participants' response times toward our experimental conditions. This approach to plotting data is particularly useful for repeated-measures analyses, as it directly illustrates the effect being tested (i.e., whether the mean difference between paired values in conditions A and B differs from 0).

Paired difference scores for each participant are represented in the graphs as coloured dots. These scores were calculated by subtracting a participants' mean reaction time for one condition of interest (e.g., Experiment 1's inconsistent pseudo-suffix condition) from that of another experimental condition (e.g., Experiment 1's consistent pseudo-suffix condition). The figures also show the mean paired difference score for each plotted comparison (represented with diamonds) and include boxplots to show the scores' distributions.

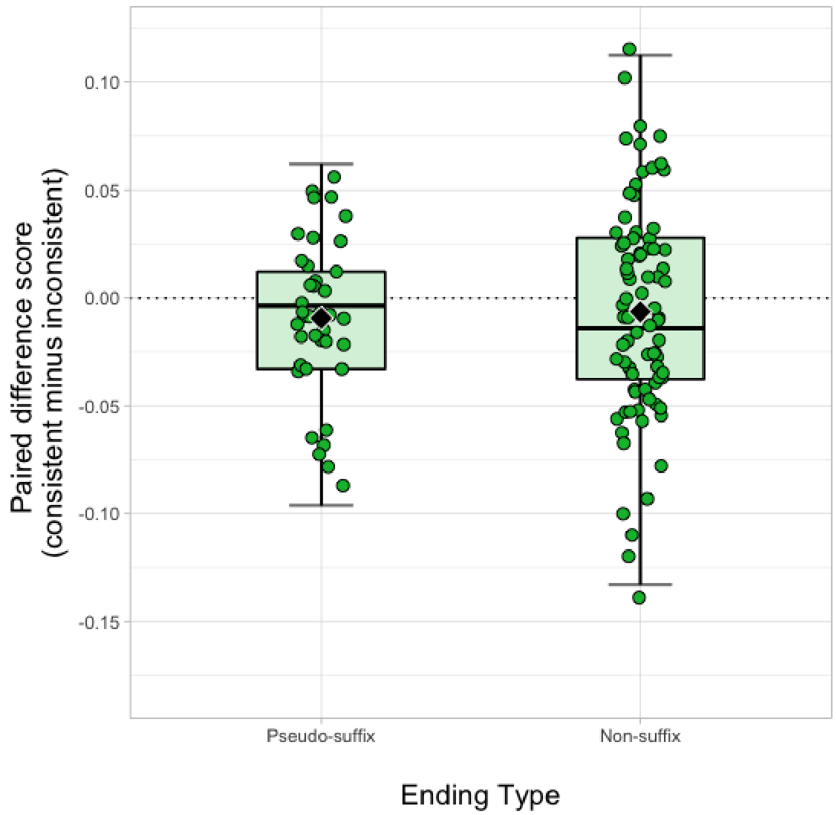


Figure 2.1. Paired difference scores showing Experiment 1's effect of endings' stress consistency on lexical decision response times, presented as a function of ending type. Scores were calculated by subtracting each participant's mean response time in the inconsistent conditions from their mean response time in the corresponding consistent conditions. As such, negative scores indicate that participants were quicker to respond to consistent items than inconsistent ones; positive scores indicate that they were quicker to respond to inconsistent items than consistent ones. Scores at 0 indicate no difference in response times across conditions. Diamonds represent the mean paired difference for each comparison.

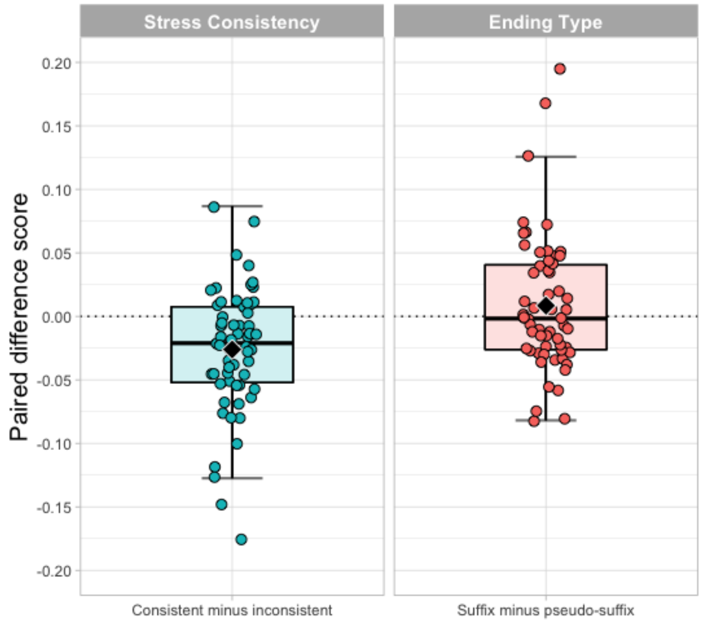


Figure 2.2. Paired difference scores showing Experiment 2's effects of endings' stress consistency (left) and ending type (right) on lexical decision response times. Stress consistency scores were calculated by subtracting each participant's mean response time in the inconsistent condition from their mean response time in the consistent condition. Negative scores indicate that participants were quicker to respond to consistent than inconsistent items. Ending type scores were calculated by subtracting each participant's mean response time in the pseudo-suffix condition from their mean response time in the suffix condition. Negative scores would indicate that participants were quicker to respond to suffixed than pseudo-suffixed items. Scores at 0 indicate no difference in response times across conditions. Diamonds represent the mean paired difference for each comparison.

2.8 APPENDIX C: EXPERIMENT 2 ITEMS

A list of the real word items used in Experiment 2, presented by condition.

SUFFIX	PSEUDO-SUFFIX		SUFFIX CHARACTERISTICS (Fudge, 1984)
Consistent	**Consistent**	**Inconsistent**	
beggar *liar*	*pillar* *mortar*	*cigar* *guitar*	Agent noun-forming; stress-neutral
herbal *tidal*	*metal* *loyal*	*cabal* *banal*	Adjective-forming; pre-stressed
corkage *outage*	*cabbage* *adage*	*garage* *montage*	Noun-forming; stress-neutral
service *justice*	*office* *lattice*	*suffice* *police*	Noun-forming; pre-stressed
heiress *actress*	*access* *mattress*	*caress* *duress*	Feminine noun-forming; stress-neutral
movement *payment*	*garment* *pigment*	*cement* *lament*	Noun-forming; stress-neutral
failure *seizure*	*future* *stature*	*mature* *brochure*	Result of an action; stress-neutral
deeply *purely*	*rally* *belly*	*supply* *imply*	Adverb-forming; stress-neutral
foolish	**nourish*		Adjective-forming; stress-neutral
rhythmic	**plastic*		Adjective-forming; pre-stressed
tension	**version*		Result of an action or process; pre-stressed
farmer	**soccer*		Agent noun-forming; stress-neutral
joyous	**viscous*		Adjective-forming; stress-neutral
ripen	**raven*		Verb-forming; stress-neutral
fearsome	**gruesome*		Adjective-forming; stress-neutral
freedom	**random*		Noun-forming; stress-neutral

Notes. Items marked with asterisks were excluded from the analysis of endings' stress consistency (pseudo-suffix consistent vs. pseudo-suffix inconsistent). Pre-stressed suffixes assign stress to a syllable before the one containing the suffix. Stress-neutral suffixes do not affect the stress patterns of the words to which they attach.

CHAPTER 3
GETTING TO KNOW THE STRESS NEIGHBOURHOOD: STATISTICAL LEARNING OF WORD ENDINGS AS STRESS CUES IN ENGLISH

3.1 ABSTRACT

Despite evidence that readers are sensitive to statistical associations between written word endings and lexical stress in English, the mechanism behind that sensitivity is unclear. In the current study, we test predictions derived from statistical learning principles to explore the claim that readers' exposure to written stress cues drives their use of those cues during visual word recognition. English-speaking adults completed a visual lexical decision task; items differed in the strength and prevalence of their ending's association with stress in English texts (quantified with a word's proportion of stress friends, number of stress friends, and number of stress enemies). After accounting for proportion of stress friends, number of stress friends had no effect on lexical decision latencies. However, after accounting for number of stress friends, participants were quicker to recognize words with larger than smaller proportions of stress friends. Separate analyses showed that this effect emerges from two competing sources: having many stress friends facilitated quicker lexical decisions for words, while having many stress enemies yielded slower lexical decisions. Together, these findings shed new light on readers' processing of lexical stress in written words.

3.2 INTRODUCTION

The vast majority of words in English have more than one syllable (Baayen, Piepenbrock, & van Rijn, 1995), and as such, multisyllabic words are crucial to a full understanding of the processes involved in word reading and visual word recognition. Traditionally, though, empirical and theoretical accounts of word reading have focused heavily on monosyllabic words (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Seidenberg & McClelland, 1989). To address this imbalance, recent attention has turned to the factors that affect multisyllabic word reading (e.g., Perry, Ziegler, & Zorzi, 2010; Yap & Balota, 2009). One such factor is lexical stress, which refers to the pattern of emphasis across syllables in a word. Evidence suggests that readers use written word endings as a probabilistic cue to stress placement, drawing on statistical associations that exist between ending spellings and stress in English (e.g., Arciuli & Cupples, 2006; Arciuli, Monaghan, & Ševa, 2010; Kelly, Morris, & Verrekia, 1998). When interpreting this work, some have suggested that statistical learning is the mechanism behind readers' sensitivity to the associations between word endings and lexical stress (e.g., Colombo, Deguchi, & Boureux, 2014). Statistical learning is a domain-general mechanism through which individuals gradually and incidentally learn of regularities in the input to which they are exposed (e.g., Perruchet & Pacton, 2006) and use their knowledge of these regularities to facilitate later processing (e.g., Arciuli & von Koss Torkildsen, 2012). The possibility that statistical learning drives readers' sensitivity to orthographic stress cues raises interesting, but untested, predictions about the impact of those cues on visual word recognition. In the current study, we provide empirical tests of these predictions.

3.2.1 WORD ENDINGS AS CUES TO LEXICAL STRESS

Statistical learning relies on the presence of regularities in the input to which people are exposed. In English, written word endings act as one such regularity. Corpus analyses have shown that many ending spellings—defined as the rime of a word's final syllable—are probabilistically associated with lexical stress patterns in disyllabic words (e.g., Arciuli & Cupples, 2006; Monaghan, Arciuli, & Ševa, 2016). For example, 91% of the disyllabic English words that end in *–el* have first-syllable stress, and as such, the ending sequence *–el* acts as a reasonably strong sublexical cue to the stress pattern of written words in which it appears (e.g., *novel*). Notably, these associations exist in the texts read by adults and those read by young children (Arciuli & Cupples, 2006; Arciuli et al., 2010). As a result, by adulthood, readers have had years of exposure to the statistical links between word endings and lexical stress in English.

Drawing on these associations, written word endings can be used to organize words into stress neighbourhoods—a term used to describe groups of words that share the same orthographic ending sequence (Colombo, 1992). As an example, the words *novel*, *label*, and *hotel* are members of the same stress neighbourhood because they share the ending *–el*. Within a neighbourhood, words that have both the same ending and same stress pattern are known as stress friends (e.g., *NOVel* and *LABel*), whereas those that have the same ending but have different stress patterns are stress enemies (e.g., *NOVel* and *hoTEL*; Burani & Arduino, 2004). These classifications allow us to quantify the extent to which a word's ending cues its stress pattern by determining a word's proportion of stress friends (calculated by dividing the word's number of stress friends by the total number of words in its stress neighbourhood). In effect, words with a large proportion of stress friends are those whose endings serve as reliable cues to stress (e.g., *NOVel*). By contrast, words with a small proportion of stress friends have endings that offer misleading stress cues (e.g., *hoTEL*). The variable serves as an index of the statistical association between a given word's ending sequence and its stress pattern, and thus represents a feature of the English writing system that might be acquired through statistical learning.

The behavioural research to date suggests that adults have indeed learned of the links between ending spellings and lexical stress, and their sensitivity to these links affects their processing of multisyllabic written words. For example, when presented with disyllabic written pseudowords and asked to select the part of the word that they would give emphasis, English-speaking adults preferentially choose the stress pattern most strongly associated with the pseudoword's ending (Arciuli & Cupples, 2006). Similar findings have been reported in Italian-speaking adults, who tend to pronounce written pseudowords with stress patterns that are consistent with their stress neighbourhoods (e.g., Colombo et al., 2014; Sulpizio, Arduino, Paizi, & Burani, 2013). Word endings' associations with stress also affect readers' naming and recognition of familiar words. For example, research in English (e.g., Kelly et al., 1998) and in Italian (e.g., Burani & Arduino, 2004; Colombo & Sulpizio, 2015) has shown that adults name written words more quickly and accurately when those words have a large proportion of stress friends than when they have a small proportion of stress friends. Similarly, English-speaking adults are quicker to make lexical decisions toward words with a large proportion of stress friends than those with a small proportion of stress friends (Kelly et al., 1998; Mundy & Carroll, 2013). These effects of word endings on the processing of known words are particularly interesting, as they suggest that readers use word endings as sublexical cues to stress even when they could conceivably rely on lexical knowledge of a word's stress pattern.

3.2.2 STATISTICAL LEARNING OF STRESS NEIGHBOURHOODS

The empirical findings to date offer convincing evidence that skilled readers of English are sensitive to the properties of stress neighbourhoods, drawing on word endings as probabilistic cues to stress when they encounter written words. However, questions remain about the mechanism by which this sensitivity emerges. As we noted above, statistical learning has been put forward as one plausible mechanism. It has been invoked in empirical studies of readers' sensitivity to stress neighbourhoods (e.g., Colombo et al., 2014) and has been implemented in connectionist models of stress placement in disyllabic word reading (Arciuli et al., 2010; Perry et al., 2010; Ševa, Monaghan, & Arciuli, 2009). According to this account, readers learn of the regularities that exist in written language through ongoing exposure to text (e.g., Arciuli & von Koss Torkildsen, 2012; Perruchet & Pacton, 2006). Crucially, for our purposes, statistical learning is an accumulative process—each new encounter with a regularity is integrated with the information gleaned from one's prior experience, modifying the strength of the statistical links that capture one's sensitivity to the underlying regularity (Thiessen & Erickson, 2013). This feature of statistical learning suggests that readers' sensitivity to a given regularity in text will vary as a function of their exposure to the written words that demonstrate it. When applied to stress neighbourhoods, this exposure translates to repeated encounters with words that share an ending sequence. Through these encounters, readers gradually extract broader statistical regularities from text (e.g., most words ending in *–el* have first-syllable stress). Ultimately, this process yields knowledge of stress neighbourhood regularities, which affects readers' stress placement when they encounter novel words (e.g., *condel*; Arciuli & Cupples, 2006) and affects their recognition of familiar words (e.g., *bushel*; Kelly et al., 1998).

To date, support for the claim that statistical learning is a mechanism behind readers' sensitivity to stress neighbourhoods has come primarily from cross-sectional research with elementary school-aged children (Arciuli et al., 2010; Colombo et al., 2014; Sulpizio & Colombo, 2013). For example, a study of English-speaking 5- to 12-year-olds found that older children made greater use of word endings to place stress in written pseudowords than did younger children (Arciuli et al., 2010). This developmental pattern aligns with similar trends seen in studies with Italian-speaking children, where the effect of word endings on stress placement is slightly larger among children in Grade 4 than children in Grade 2 (Colombo et al., 2014; Sulpizio & Colombo, 2013). These cross-sectional findings are certainly compatible with a statistical learning mechanism—older children have had more exposure to text than their younger peers, affording them more opportunities to learn of the various regularities that are present in text. As a result, we would indeed expect children to become more sensitive to—and make greater use of—stress neighbourhood regularities as age increases. While compelling, the cross-sectional findings described above provide only indirect support for the claim that statistical learning drives readers' sensitivity to stress neighbourhoods. Further empirical work is therefore needed to test the claim more directly.

In the current study, we address this gap by exploring two key questions derived from statistical learning theory. The first of these questions is based on the observation that readers gain more exposure to some stress neighbourhoods than others: some word endings occur in many written words, whereas other endings are comparatively rare. We explore whether these differences in stress neighbourhood size affect readers' multisyllabic word recognition—a possibility that has been overlooked in the prior research that focused solely on a word's proportion of stress friends (e.g., Kelly et al., 1998). The second question explored in our study addresses the fact that stress neighbourhoods are composed of two elements—stress friends and stress enemies—both of which might plausibly contribute to readers' sensitivity to a word ending's link with stress. However, we do not yet have data on the respective effects of stress friends and stress enemies on adults' recognition of written words. In the sections that follow, we elaborate on these research questions in turn.

3.2.3 DIFFERENCES IN READERS' EXPOSURE TO STRESS NEIGHBOURHOODS

In the behavioural research described above, stress neighbourhoods were characterized by their proportion of stress friends (e.g., Arciuli & Cupples, 2006; Burani & Arduino, 2004; Kelly et al., 1998). This is certainly a useful approach, as proportion of stress friends gives a key measure of the extent to which a word's ending cues its stress pattern. However, it cannot capture differences in the extent of readers' exposure to stress neighbourhoods. This is a noteworthy limitation to the variable, particularly given that exposure to regularities is central to statistical learning. Thus, in order to test the claim that statistical learning is the mechanism behind readers' sensitivity to stress neighbourhoods, we need to consider differences in words' proportion of stress friends alongside differences in their stress neighbourhood size. If readers learn of stress neighbourhood regularities through statistical learning, we would expect them to be particularly sensitive to regularities that are widely reflected in text and less sensitive to regularities that are uncommon.

Some examples may help to illustrate why this distinction is important when exploring the claim that statistical learning drives readers' sensitivity to word endings as stress cues. The words *NOVel* and *ballOON*, for instance, have similar proportions of stress friends (91% and 94%, respectively). However, *–el* is by far the more common of the two endings; it occurs in 166 words, whereas *–oon* occurs in only 34 words. As a result, the word *NOVel* has a fairly large number of stress friends, and by contrast, *ballOON* has relatively few (150 and 32 stress friends, respectively). Readers therefore encounter far more examples of the association between *–el* and first-syllable stress than examples of the association between *–oon* and second-syllable stress. If successive exposures strengthen one's implicit sensitivity to a stress cue, we would expect a word's number of stress friends to affect readers' processing of multisyllabic words, even after taking into account the effect of that word's proportion of stress friends.[7]

One existing study offers preliminary support for the possibility that both proportion and number of stress friends are relevant to readers' processing of multisyllabic written words. In that study, Sulpizio and colleagues (2013) had adult readers of Italian complete a pseudoword naming task. All items had ending sequences that were associated with either penultimate stress (*binTOro*) or antepenultimate stress (*BIRnolo*), and the researchers separately calculated each pseudoword's proportion of stress friends and absolute number of stress friends from among real words with those endings in Italian. In post hoc analyses, they explored the rate at which readers produced the stress pattern most strongly associated with each pseudoword's ending and found that both proportion and number of stress friends predicted stress assignment. That is, the larger a pseudoword's proportion of stress friends and the greater its number of stress friends, the more likely participants were to pronounce it with the stress pattern suggested

[7] For the sake of consistency with prior work by Sulpizio, Arduino, Paizi, and Burani (2013), we focus on number of stress friends as a proxy for stress neighbourhood size. However, note that the results and conclusions that we report for research question #1 do not change when we analyze our data using a word's total number of stress neighbours in place of number of stress friends.

by its ending. Interestingly, these analyses showed that the effect of number of stress friends on stress placement was stronger than that of proportion. Based on this finding, the study's authors concluded that "to drive stress assignment, a[n ending] sequence not only has to be associated with a given stress pattern in a relatively high proportion of cases but also has to appear in a large number of stress neighbours" (Sulpizio et al., 2013, p. 61).

This interpretation is consistent with a key prediction derived from the statistical learning framework: that the effect of stress neighbourhoods on readers' processing of written words should be strongest when the associations between word endings and lexical stress are widely encountered. To our knowledge, though, Sulpizio and colleagues' (2013) findings—which were based on post-hoc analyses—are the only reported test of this claim. And importantly, their findings emerged from analyses of pseudoword naming, which required the use of a restricted range of proportion of stress friends (65% to 100%, rather than the full range from 0% to 100%). This limited range of proportions was appropriate for their purposes, as it was the only meaningful way to establish a "correct" stress pattern for their pseudoword items. Even so, restricting one of the central variables in their analysis is likely to have influenced their results. To address this issue in the current study, we examine the respective effects of a word's proportion and number of stress friends on adults' recognition of real English words. Doing so allows us to include the full range of possible proportion of stress friends scores, which offers a stringent test of the claim that a word ending's effect on stress placement depends on both the strength of the ending's association with stress and the number of words demonstrating that association.

Our key prediction is that after controlling for a word's proportion of stress friends, adults will recognize words more quickly when they have many stress friends than when they have fewer stress friends. This finding would support the claim that readers' exposure to a regularity increases their sensitivity to, and use of, that regularity when processing written words. Conversely, though, we also expect a word's proportion of stress friends to affect word recognition after accounting for its number of stress friends. This prediction builds on the well-established evidence that a word's proportion of stress friends affects readers' naming and recognition of written words (e.g., Burani & Arduino, 2004; Colombo & Sulpizio, 2015; Kelly et al., 1998; Mundy & Carroll, 2013). If this finding is robust, it should not be subsumed by measures of stress neighbourhood size (which, when considered alone, cannot capture the strength of an ending's association with stress). By testing the respective roles of a word's proportion and number of stress friends, our study gives insight into the characteristics of stress neighbourhoods that affect readers' processing of multisyllabic written words.

3.2.4 THE COMPONENTS OF STRESS NEIGHBOURHOODS

At its core, exposure to a stress neighbourhood encompasses a reader's relative levels of exposure to words whose stress patterns align with (e.g., *NOVel*) and differ from (e.g., *hoTEL*) the stress pattern most strongly associated with that neighbourhood (e.g., first-syllable stress for the ending *–el*). In the behavioural research to date, these relative levels of exposure have been captured coarsely with a word's proportion of stress friends (e.g., Arciuli & Cupples, 2006; Kelly et al., 1998). While useful in establishing that endings serve as cues to stress in written words, this approach cannot speak to the components of a stress neighbourhood that drive word endings' effects on visual word recognition. To do so, we must consider the separate effects of a word's number of stress friends and its number of stress enemies.

Although this question has not yet been addressed empirically, its answer is key to interpretations that have been offered for prior findings on readers' sensitivity to stress neighbourhoods. For example, Kelly and colleagues' (1998) seminal research on the topic showed that English-speaking adults respond more quickly and accurately to written words whose endings were consistent with their stress patterns than words whose endings were inconsistent with their stress patterns. The authors suggested two sources for this effect: having many stress friends facilitates readers' processing of a written word, while having many stress enemies impedes processing (Kelly et al., 1998, p. 827). This interpretation has intuitive appeal, but it requires further testing—the effects that Kelly and colleagues (1998) observed could be driven solely by facilitation from stress friends, by hindrance from stress enemies, or by both (as the study's authors contended). Moreover, while the statistical learning framework suggests that exposure drives readers' sensitivity to stress neighbourhoods, it does not indicate the specific regularities that readers draw on when processing written words. Thus, it is possible that readers draw on their exposure to one component of stress neighbourhoods more strongly than the other. Without disentangling the respective effects of a word's stress friends and stress enemies, we cannot draw firm conclusions about the source of word endings' impact on written word recognition.

As such, our second goal for the current study is to determine whether number of stress friends and/or number of stress enemies influence readers' processing of multisyllabic written words. We focus specifically on the extent to which the two variables separately affect written word recognition. That is, after controlling for a word's number of stress friends, do readers recognize words with many stress enemies more slowly than words with few? Conversely, after controlling for a word's number of stress enemies, are readers quicker to recognize words that have many stress friends than those with few? Through targeted analyses that quantify the respective effects of a word's number of stress friends and stress enemies, our study will serve to characterize the nature of, and mechanism behind, readers' sensitivity to stress neighbourhoods.

3.2.5 THE CURRENT STUDY

We address two research questions in the current study, both of which seek to clarify the statistical features of stress neighbourhoods that affect readers' processing of multisyllabic English words. Our first research question compares the respective effects of a word's proportion and number of stress friends, testing two hypotheses: (a) after controlling for proportion of stress friends, adults will be quicker to identify words with many stress friends than those with fewer, and (b) after controlling for number of stress friends, adults will be quicker to identify words the greater their proportion of stress friends. The first of these hypotheses tests a key prediction derived from statistical learning principles—that the extent of readers' exposure to a regularity should affect the extent to which they learn of and use that regularity to process written words. The latter hypothesis builds on prior evidence that a word's proportion of stress friends affects the speed with which it is recognized (e.g., Kelly et al., 1998; Mundy & Carroll, 2013), aiming to establish whether this effect persists after accounting for differences in stress neighbourhood size. Our second research question disentangles the two components of a stress neighbourhood, testing the predictions that: (a) after controlling for a word's number of stress enemies, adults will recognize words with many stress friends more quickly than those with few, and (b) after controlling for number of stress friends, adults will be slower to recognize words with many stress enemies than those with few. Exploring these possibilities will help us to better understand the mechanism that drives readers' sensitivity to word endings as stress cues in English.

To address these questions, we asked English-speaking adults to complete a visual lexical decision task. This task has been successfully used in prior studies on English speakers' sensitivity to stress neighbourhoods (Kelly et al., 1998; Mundy & Carroll, 2013), and for the current study's purposes, it has two key advantages. The first is that lexical decision tasks are based around the characteristics of real words (cf. Sulpizio et al., 2013). By using real words as target items, we can take advantage of the full continuum of proportion values (from 0% to 100% stress friends, a range not possible in pseudoword tasks) for a stringent test of the claim that number of stress friends affects word recognition after controlling for proportion of stress friends. As a second advantage, using a lexical decision task avoids some of the complications involved with real-word naming tasks. For example, lexical decision does not require us to control for words' initial phonemes, as is necessary in tasks that measure naming latencies (Kessler, Treiman, & Mullennix, 2002). More notably, evidence suggests that a word's stress pattern affects naming responses, but it does not affect the speed of lexical decisions toward disyllabic words (e.g., Mundy & Carroll, 2013; Yap & Balota, 2009). Taken together, these characteristics of visual lexical decision tasks afford us the flexibility to explore our research questions without introducing challenging confounds. Importantly, although the task does not involve phonological output, phonological variables can affect lexical decision performance (e.g., Halderman, Ashby, & Perfetti, 2012; Kelly et al., 1998; Mundy & Carroll, 2013).

3.3 METHOD

3.3.1 PARTICIPANTS

Participants were 56 university students aged 17 and older (44 women; mean age = 20.97 years; *SD* = 3.01 years), all of whom spoke English as a first language and had normal or corrected-to-normal vision. Four additional participants were tested, but their data were not usable (two participants did not speak English as a first language; two participants' data were lost due to technical error). As compensation, participants earned either $5 or 0.5 percentage points towards a university course.

3.3.2 MATERIALS

Selection of word endings. We adopted the definition of word endings provided by Arciuli and Cupples (2006): the letter string beginning at the second phonemic vowel of a disyllabic word, plus any consonants that follow. Effectively, this ending unit corresponds with the rime of a word's second syllable. We selected a total of 31 word endings, which were isolated from the 14661 disyllabic English words listed in the CELEX database (excluding only the 1305 hyphenated entries, e.g., *horn-rimmed*; Baayen et al., 1995). Of our 31 word endings, 18 were primarily associated with first-syllable stress and 13 were primarily associated with second-syllable stress; the strength of these associations varied on a continuous scale.

Selection of items. As items for the lexical decision task, we selected 180 words that contained one of our chosen word endings (see Appendix D). We reviewed all items for alternate spellings and pronunciations using the CELEX English database (Baayen et al., 1995) and verified that all pronunciations of a given item held the same lexical stress pattern. For example, the word *extract* was excluded because it can be pronounced with either first- or second-syllable stress. Our final list of items included an equal number of words with first- and second-syllable stress, and included six items per ending in all but two cases (*–eive* and *–olve*, each of which had three associated items). These exceptions emerged because only three words ending in *–eive* met our selection criteria; as such, we included three items with the ending *–olve* to ensure an equal number of items with first- and second-syllable stress.

Item-level control variables. Items ranged in frequency from 1 to 100 occurrences per million (M = 24.28; SD = 24.11; Baayen et al., 1995) and in length from 4 to 10 letters (M = 6.41; SD = 1.06). We controlled for variance associated with these item characteristics in all analyses, as both are predictive of lexical decision latencies (e.g., Yap & Balota, 2009). We also controlled for variance associated with two measures of our items' orthographic similarity to other words in the lexicon: mean bigram frequency and orthographic Levenshtein distance (both obtained from the English Lexicon Project; Balota et al., 2007). Mean bigram frequency captures the average number of times each two-letter substring within a word appears in a corpus; for our items, mean bigram frequencies ranged from 584 to 3419.2 (M = 1861.06; SD = 641.19). Orthographic Levenshtein distance (OLD) is a measure of similarity that is based on the minimum number of letter substitutions, deletions, or insertions required to convert one string into another. Using this definition, OLD captures the mean distance from a word to its 20 closest orthographic neighbours (see Yarkoni, Balota, & Yap, 2008 for details). OLD scores for our items ranged from 1.7 to 4.4 (M = 2.33; SD = 0.47). When compared to traditional measures of neighbourhood size (e.g., orthographic N; Coltheart, Davelaar, Jonasson, & Besner, 1977), OLD allows for a more flexible definition of orthographic similarity and is better able to predict lexical decision performance (Yarkoni et al., 2008). It is particularly valuable for use with long and multisyllabic words, which have few orthographic neighbours under traditional neighbourhood size metrics (Coltheart et al., 1977).

Temporal control variables. In addition to the item-level variables described above, we controlled for trial-by-trial temporal dependencies in participants' lexical decision latencies. First, we controlled for trial number—the rank order in which items were presented. This order was unique to each participant, as item presentation was randomized with each administration of the lexical decision task. Inclusion of trial number as a control variable acknowledges that participants' reaction times tend to change over the course of an experimental task, reflecting the effects of learning and/or fatigue as the task progresses (Baayen & Milin, 2010). Secondly, we controlled for the participant's reaction time to the preceding trial (Preceding RT; initial trials were given the participant's mean response latency, per Baayen & Milin, 2010). During visual word recognition, evidence suggests that participants' trial-by-trial responses are partially interdependent—slower responses on one trial predict slower responses on the next, with large effects (e.g., De Vaan, Schreuder, & Baayen, 2007; Kuperman, Schreuder, Bertram, & Baayen, 2009). Given these temporal dependencies, controlling for trial number and preceding RT improves the precision of models and allows us to better estimate the contributions made by the stress neighbourhood predictors that are of theoretical interest to our study.

Calculation of stress neighbourhood variables. Our key predictors of interest quantified various aspects of a word's stress neighbourhood, capturing a word's (a) proportion of stress friends, (b) number of stress friends, and (c) number of stress enemies. We calculated each item's number of stress friends by counting all disyllabic words with the same ending and stress pattern as that item in the CELEX database (Baayen et al., 1995), then subtracting 1 to exclude the item itself. Items' number of stress friends ranged from 1 to 217 (M = 42.72; SD = 47.20). Similarly, the calculation for number of stress enemies was a simple tally of the number of disyllabic words that had the same ending sequence as an item, but had a different stress pattern. Number of stress enemies ranged from 1 to 218 (M = 21.84; SD = 36.05). Words pronounced with both first- and second-syllable stress (e.g., *extract*) were included in these calculations, and were counted once as a stress friend and once as a stress enemy. A word's proportion of stress friends was calculated by dividing the number of stress friends by the total number of disyllabic words with the same ending as the item. Proportions ranged from 0.05 to 1.0 (M = 0.69; SD = 0.29). Given the nature of these calculations, the three variables are not fully independent of each other (though they are separable; see section 3.2.3 for an example illustrating this). We took this into consideration when selecting items to avoid problematically high correlations between variables (see Table 3.1).

Table 3.1.

Zero-order correlations among all stress neighbourhood predictor variables and item-level control variables.

Measures	1.	2.	3.	4.	5.	6.
1. Proportion of stress friends	—					
2. Number of stress friends	$.43^{***}$	—				
3. Number of stress enemies	$-.67^{***}$	-.07	—			
4. Word frequency	-.10	$.14^{\dagger}$	.03	—		
5. Word length	.08	$-.31^{***}$	$-.21^{**}$	.08	—	
6. OLD	$.14^{\dagger}$	$-.29^{***}$	$-.24^{**}$	-.06	$.76^{***}$	—
7. Mean bigram frequency	$-.18^{*}$	-.04	$.18^{*}$	.01	.04	$-.27^{***}$

Notes. Significant and marginal correlations are noted as follows: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$, $^{\dagger}p < .10$. OLD = Orthographic Levenshtein Distance.

Creation of pseudowords. As foils for our lexical decision task, we created 180 pseudowords by changing one to two consonants in the initial part of each real-word target item (e.g., Ford, Davis, & Marslen-Wilson, 2010). These pseudowords preserved the word endings of their real-word counterparts, maintained the same number of letters and phonemes, and obeyed the orthographic and phonotactic constraints of English. We constructed the pseudowords with the intent that participants would read them as disyllabic; this was done to minimize participants' ability to systematically distinguish words from pseudowords based on number of syllables (cf. Kelly et al., 1998). Together, this process resulted in pseudowords that were closely matched to the real-word items used in our lexical decision task.

3.3.3 PROCEDURE

Participants were tested individually on a PC laptop running DirectRT software (Jarvis, 2008b). Each trial began with a fixation cross presented for 1000ms at the centre of a white screen. Next, a word or pseudoword item replaced the fixation cross (Times New Roman font, size 40). Participants were asked to decide whether the target string was a real English word, and to respond by pressing a green “Y” button if the item was a real word or a red “N” button if it was not. The position of the response keys matched the participant’s handedness such that each participant indicated real words with their dominant index finger. On each trial, participants were given up to 5000ms to respond, at which point the target disappeared and the screen went blank. Participants completed eight practice trials, during which time the experimenter provided corrections for miscategorised words. They then completed the 360 experimental trials, which were presented in a randomized order without feedback. The lexical decision task took approximately 20 minutes to complete.

3.4 RESULTS

Data were analyzed using linear mixed-effects models with participants and items as crossed random effects (Baayen, Davidson, & Bates, 2008). The analyses were conducted in R (R Core Team, 2017) using the lme4 package’s *lmer* function (Bates, Maechler, Bolker, & Walker, 2017).

3.4.1 DATA CLEANING

Prior to analysis, we excluded 10 items whose mean accuracy rates were lower than 50% across participants (e.g., Jiang, 2012). As would be expected, the excluded words were of low frequency (e.g., *luddite*, *subsume*; frequencies of 1 to 3 per million), and participants’ low accuracy rates indicate that they recognized these items at a below-chance level. Across participants, accuracy for each of the remaining 170 items was at ceiling (see Appendix D; overall accuracy across participants and items: M = 92.6%; SD

= 3.6%). As such, the analyses reported below use reaction times to correctly identified words as their outcome measure. Reaction times greater than 5000ms were excluded as extreme values (3 data points; 0.03% of all data), as this was the maximum length of time that items were displayed. We also excluded reaction times that were more than 2.5 standard deviations above each participant's mean reaction time (cutoff selected based on Van Selst & Jolicoeur, 1994; non-recursive moving criterion method). Excluding these outliers resulted in a loss of 244 data points (2.6% of all data). There were no overly short reaction times requiring exclusion.

Following these outlier exclusions, we transformed variables that exhibited positive skew (Tabachnick & Fidell, 2007). Our outcome variable (reaction time to correctly identified words) was logarithmically transformed, as were several of our predictor variables (word frequency, orthographic Levenshtein distance, mean bigram frequency, number of stress friends, number of stress enemies, and reaction time to the preceding trial). Word length was square-root transformed. These transformations successfully addressed the slight heteroscedasticity that was present in the residuals of models computed using raw scores. All predictor variables were scaled and centered prior to analysis.

3.4.2 PROCEDURES FOR MODEL CONSTRUCTION AND COMPARISON

Model comparisons. To address our research questions, we constructed two sets of models that include differing combinations of our variables of interest (Models 1–4 explore our first research question; Models 5–8 explore our second). Because of the challenges inherent in estimating *p*-values for mixed-effects models (e.g., Baayen et al., 2008), we structured our models to allow for two methods of evaluating statistical significance. First, we provide fixed effect estimates with 95% Wald confidence intervals in Figure 3.1 (addressing research question 1) and Figure 3.3 (addressing research question 2). An advantage to this approach is that it lets us interpret the significance of individual fixed effects—those with confidence intervals that exclude zero can be taken as significant predictors in a model. However, some evidence suggests that the use of Wald values yields a higher Type I error rate than that seen in other methods of assessing statistical significance in mixed-effects models (including likelihood-ratio tests; Luke, 2017).

To address this issue with Wald values, we also evaluated statistical significance by comparing models with likelihood-ratio tests. Simulations suggest that likelihood-ratio tests are an appropriate way to generate *p*-values for mixed-effects analyses of typically-sized psycholinguistic datasets (Barr, Levy, Scheepers, & Tily, 2013). All models being compared had the same outcome variable (log-transformed reaction time to correctly identified words) and had the same random effects, but differed in the number of fixed effects included as predictor variables. This approach yields nested models, which allows us to compare two models that differ only in their inclusion or exclusion of a predictor variable of interest.[8] For example, we might compare a model containing the fixed effects of only our control variables to a model containing those of the controls plus proportion of stress friends (Models 1 and 2, as shown in Figure 3.1). Broadly, a difference in the likelihood of these models can be attributed to the proportion of stress friends fixed

[8] Conceptually, these nested model comparisons function much like a hierarchical ordinary least squares regression analysis. For example, comparing the base model with Model 1 (both are described below) is analogous to including all control variables in Step 1 of a regression model. Comparing Models 1 and 2 is then like adding a word's proportion of stress friends in Step 2 of that regression, and comparing Models 2 and 4 adds number of stress friends in Step 3.

effect, as its inclusion/exclusion is the sole difference between the two models. Strictly speaking, however, likelihood-ratio tests do not test for the significance of individual fixed effects. Instead, these tests compare the two nested models in aggregate. With this in mind, we use the likelihood-ratio tests reported below as a converging source of evidence for variables' statistical significance, to be interpreted alongside the confidence intervals in Figures 3.1 and 3.3.

Fixed effects. The fixed effects included in our models were either properties of the items used in our lexical decision task or were measures of trial-by-trial dependency. All were continuous variables. As controls, we included the fixed effects of word frequency, word length, orthographic Levenshtein distance, mean bigram frequency, trial number, and preceding RT. Models 1 and 5 included only these fixed effects; they served as control models for the analyses addressing our first and second research questions, respectively. All other models included these same control variables, plus one or more of the stress neighbourhood predictor variables: proportion of stress friends, number of stress friends, and number of stress enemies (see Figures 3.1 and 3.3). The particulars of these models are described in more detail below. Our focus in the analyses that follow is on whether, and to what extent, the fixed effects of these stress neighbourhood variables predict lexical decision latencies.

Random effects. To establish our models' random-effects structures, we used an iterative model selection process recommended by Bates and colleagues (2015). We began by attempting to fit a maximal random-effects structure to the most complex model in a given set (i.e., Model 4 for our first research question and Model 8 for our second research question; these models contained the largest number of fixed effects). To address convergence issues, we refit both models by excluding the correlation parameters between random slopes and intercepts. We then conducted principal components analyses (PCAs) on the resulting zero-correlation-parameter models' random-effects covariance matrices (RePsychLing package; Baayen, Bates, Kliegl, & Vasishth, 2015). These analyses suggested that both models were over-parameterized (i.e., they yielded principal components accounting for zero variance; Bates et al., 2015). As such, we dropped the weakest random effects from each model and repeated the PCAs until we reached the smallest the number of principal components that accounted for a cumulative 100% of variance. Bates and colleagues (2015) argue that these parsimonious random-effects structures yield models that are supported by their data and that protect against Type I error to roughly the same extent as a maximal random-effects model (cf. Barr et al., 2013).

Through this model selection process, we arrived at simplified random-effects structures for Models 4 and 8. We then used these same random-effects structures for the nested models to which Models 4 and 8 were being compared. Specifically, we used the random-effects structure of Model 4 for Models 1, 2, and 3 (addressing our first research question), and similarly, we used the random effects structure of Model 8 for Models 5, 6, and 7 (addressing our second research question). We used the process described above to confirm that these random effects were appropriate for all models. Our models' final random-effects structures are summarized in Appendix E.

Multicollinearity. Given the potential for our predictor variables of interest (in particular, proportion of stress friends, number of stress friends, and number of stress enemies) to be highly correlated, we inspected our models for multicollinearity. Across all of the models that we report below, variance inflation factors were smaller than 3 and kappa values were smaller than 4. These indices are within the acceptable range for the use of mixed-effects models (Baayen, 2008), and they suggest that multicollinearity is not an issue in our models.

Calculating percent reduction in item variance. To establish the explanatory value of our predictor variables, we constructed a base model containing only the crossed random effects of participant and item. This base model provides us with an estimate of the variability in reaction time that is attributable to participants ($SD = 0.059$) and items ($SD = 0.034$), respectively (residual $SD = 0.080$). For our purposes, the key function of the base model is establishing a baseline item random effect. Adding our predictor variables as fixed effects in subsequent models provides information about differences between items (e.g., their word frequencies; their proportion of stress friends, etc.). When that information is relevant to observed differences in reaction time, some of the variance associated with items is explained, and the random effect of items is reduced. Thus, in the model comparisons below, we report percent reductions in item variance (calculated per Baayen, 2008) as an index of the extent to which each item-specific predictor explains variability associated with the item random effect.

3.4.3 RESEARCH QUESTION 1: EFFECTS OF PROPORTION AND NUMBER OF STRESS FRIENDS

Our first research question concerns the respective effects of a word's proportion and number of stress friends on participants' lexical decision latencies. Specifically, we tested two hypotheses: (a) after controlling for a word's proportion of stress friends, participants would respond more quickly to words with more stress friends than those with fewer, and (b) after controlling for a word's number of stress friends, participants would respond more quickly to words with a larger proportion of stress friends than those with a smaller proportion of stress friends.

To that end, we created four models predicting participants' log-transformed lexical decision latencies. First was a control model (Model 1), which included word frequency, word length, orthographic Levenshtein distance, mean bigram frequency, trial number, and preceding RT as fixed effects. All other models contained this same set of control variables. Specifically, Model 2 included the control variables and proportion of stress friends, but excluded number of stress friends. Similarly, Model 3 included the controls and number of stress friends, but excluded proportion of stress friends. Finally, we created a full model (Model 4), which included the controls and both proportion and number of stress friends. Figure 3.1 shows the standardized estimates and confidence intervals for each included fixed effect.

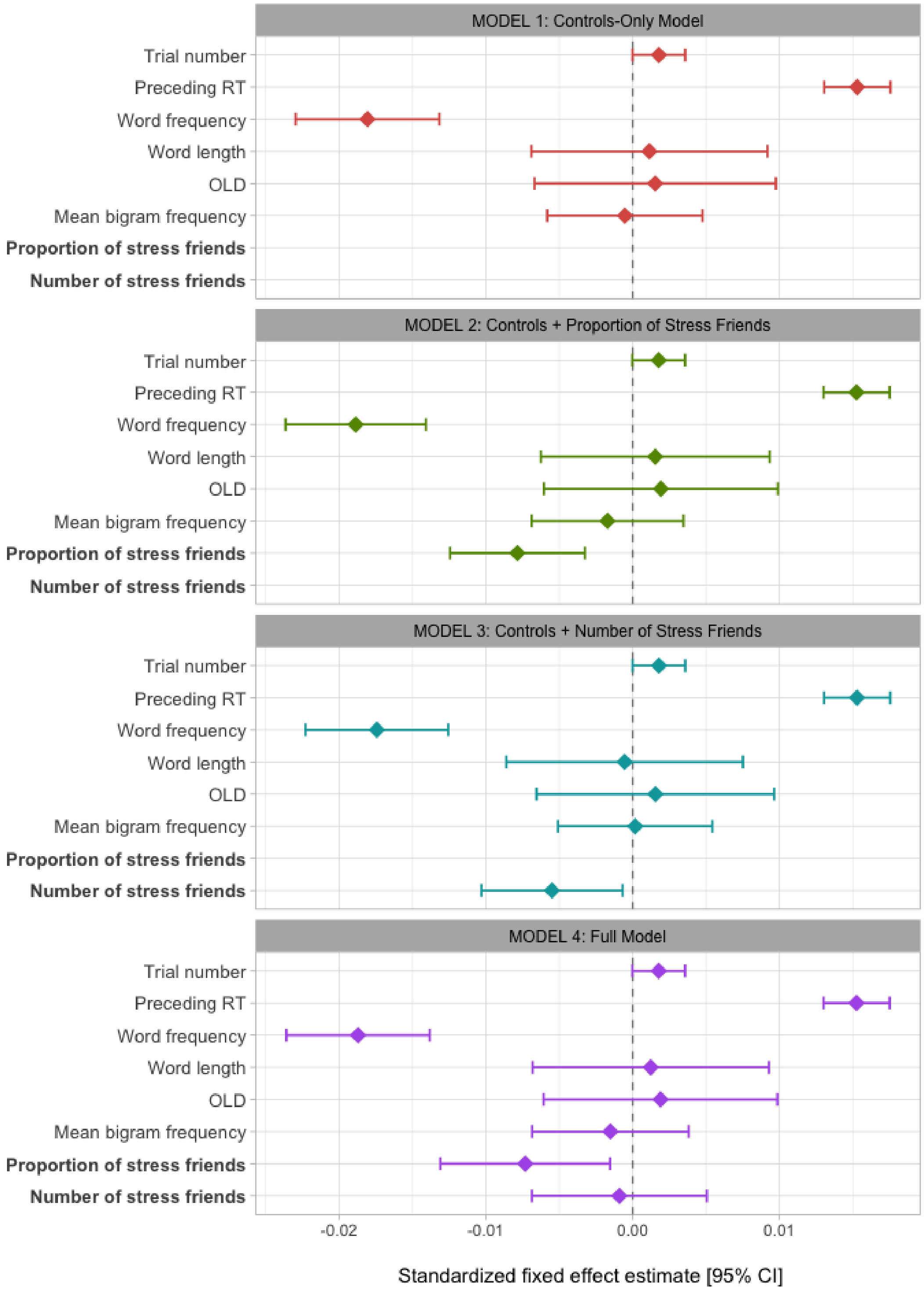


Figure 3.1. Standardized fixed effects estimates for Models 1 through 4, predicting participants' log-transformed reaction times to correctly identified words. Error bars show Wald 95% confidence intervals.

To quantify the effect of our control variables on participants' lexical decision latencies, we compared Model 1 with the intercept-only base model described above. Doing so showed that our control variables' fixed effects and random slopes contributed significantly to Model 1's goodness of fit, $\chi^2(11) = 293.20$, $p < .001$, and reduced item variance by 16.51% and participant variance by 12.81%. As Figure 3.1 indicates, some of the individual control variables (i.e., word length, OLD, and mean bigram frequency) made no contribution to Model 1. However, in light of their theoretical relevance to visual word recognition, we retained these variables in our models. The procedure through which we established our models' random-effects structures (described above) suggests that our models are not over-parameterized despite the inclusion of these nonsignificant fixed effects.

Beyond the effect of these control variables, proportion and number of stress friends (when considered independently of each other) further predicted response times. That is, comparing Models 1 and 2 showed that the fixed effect of proportion of stress friends contributed significantly to Model 2's goodness of fit, $\chi^2(1) = 10.90$, $p < .001$ (3.59% reduction in item variance), suggesting that it accounts for variance in response times beyond that accounted for by the controls. The effect estimate for proportion of stress friends was negative (see Figure 3.1)—the larger a word's proportion of stress friends, the quicker participants were to respond. Similarly, comparing Models 1 and 3 showed that the fixed effect of number of stress friends contributed significantly to Model 3's fit, $\chi^2(1) = 4.94$, $p = .03$ (1.66% reduction in item variance). Again, this effect was negative—the more stress friends a word had, the quicker participants' responses. These comparisons suggest that proportion and number of stress friends each separately predict lexical decision latencies. However, they do not address our research question, which concerns the unique effects of proportion and number of stress friends when the two variables are considered together. We turn to model comparisons that address this question next.

First, we compared Model 2 (which includes proportion of stress friends but not number) with Model 4 (which includes both proportion and number of stress friends). Effectively, this comparison tests whether a word's number of stress friends predicts lexical decision latencies after accounting for its proportion of stress friends. Contrary to our expectations, the likelihood-ratio test showed that number of stress friends had no impact on Model 4's goodness of fit (vs. that of Model 2), $\chi^2(1) = 0.09$, $p = .77$ (0.03% reduction in item variance). In other words, once we controlled for the effect of proportion of stress friends (as well as all other item-level and temporal controls), participants responded to words with few stress friends just as quickly as they did words with many stress friends.

Next, we compared Model 3 (which includes number of stress friends but not proportion) with Model 4 (which includes both proportion and number of stress friends) to establish whether proportion of stress friends contributes to lexical decision latencies after accounting for number of stress friends. The likelihood-ratio test suggests that it does—proportion of stress friends contributed significantly to Model 4's goodness of fit (vs. that of Model 3: $\chi^2(1) = 6.05$, $p = .01$; 2.00% reduction in item variance). As the bottom panel of Figure 3.1 shows, after accounting for number of stress friends (along with our other item-level and temporal controls), participants were quicker to identify words the larger their proportions of stress friends.

To further illustrate the separate effects of a word's proportion and number of stress friends on lexical decision latencies, Figure 3.2 plots their respective marginal effects as predictors in Model 4. These plots estimate change in response times across the variables' ranges, with all other predictors in Model 4 held constant at their means. Taken together, our analyses suggest that stress neighbourhood consistency (as quantified by a word's proportion of stress friends) has a reasonably robust effect on lexical decision latencies that applies across stress neighbourhoods of varying size. To further unpack this effect, our second research question explores the two components of stress neighbourhood consistency: stress friends and stress enemies.

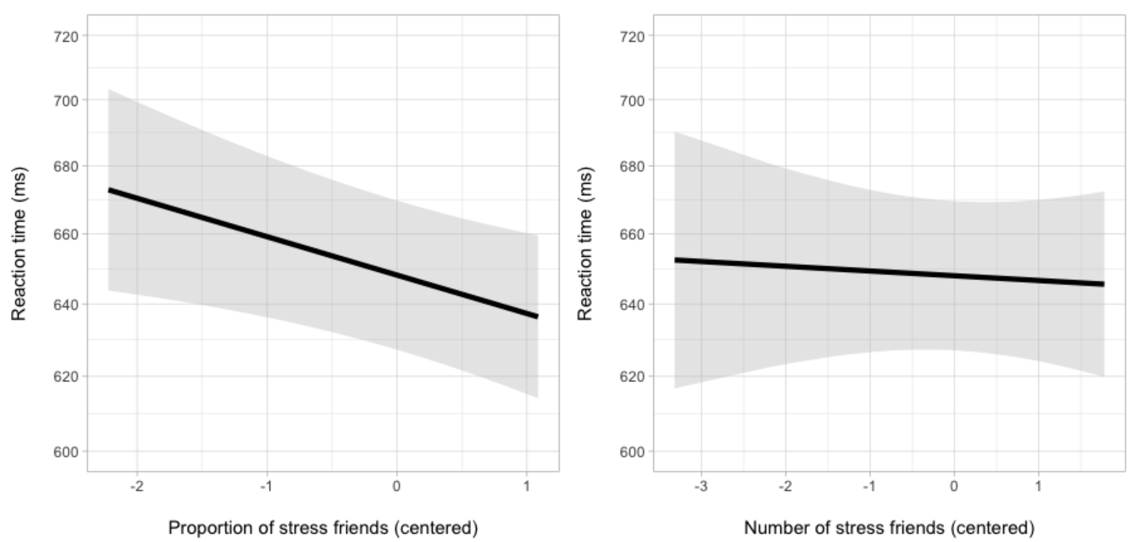


Figure 3.2. Marginal effects [95% CI] of a word's proportion of stress friends and number of stress friends, respectively, as predictors of lexical decision latency in Model 4. All other predictors in Model 4 are set to their means. Graphs' *y*-axes are shown with log-transformed coordinates.

3.4.4 RESEARCH QUESTION 2: EFFECTS OF NUMBER OF STRESS FRIENDS AND NUMBER OF STRESS ENEMIES

Our second research question concerns the respective effects of a word's number of stress friends and number of stress enemies on lexical decision latencies. Specifically, we explored the possibilities that: (a) the larger a word's proportion of stress friends, the quicker participants' lexical decisions, and conversely, (b) the larger a word's number of stress enemies, the slower participants' lexical decisions. In both cases, we were most interested in whether number of stress friends and enemies would uniquely predict response times after controlling for the effect of the other.

For this research question, we created four new models; as previously, all models predicted participants' log-transformed lexical decision latencies. First was a control model (Model 5), which included word frequency, word length, orthographic Levenshtein distance, mean bigram frequency, trial number, and preceding RT as fixed effects. Note that although our two control models (Models 1 and 5) are very similar, they have slightly different random-effects structures (see Appendix E), reflecting the different predictors used to address our two research questions. As before, all subsequent models contained the control variables present in Model 5. Model 6 included the control variables and number of stress friends, but excluded number of stress enemies. Model 7 included the controls and number of stress enemies, but excluded number of stress friends. Finally, the full model (Model 8) included the control variables along with both number of stress friends and number of stress enemies. Figure 3.3 shows standardized estimates and confidence intervals for each included fixed effect.

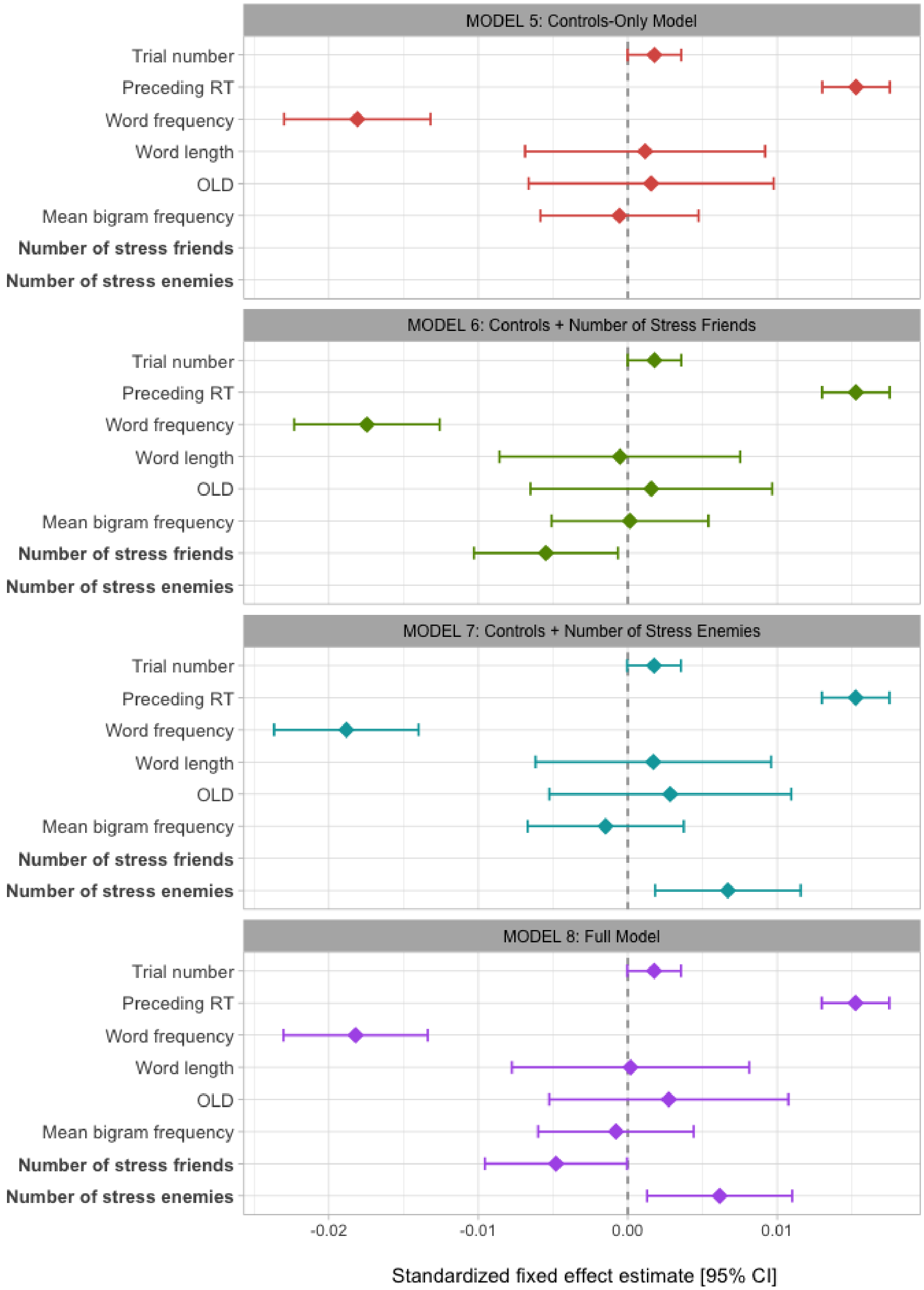


Figure 3.3. Standardized fixed effects estimates for Models 5 through 8, predicting participants' log-transformed reaction times to correctly identified words. Error bars show Wald 95% confidence intervals.

As with our first research question, we began by quantifying the effect of our control variables on participants' lexical decision latencies. We did so by comparing Model 5 with the intercept-only base model. Unsurprisingly, the results were highly similar to those reported earlier with Model 1. Our control variables' fixed effects and random slopes contributed significantly to Model 5's goodness of fit, $\chi^2(11) = 294.38$, $p < .001$, and reduced item variance by 16.57% and participant variance by 12.77%. We then explored whether number of stress friends and enemies would predict participants' response latencies (when considered independently of each other). Indeed, comparing Models 5 and 6 showed that the fixed effect of number of stress friends contributed significantly to Model 6's goodness of fit, $\chi^2(1) = 4.91$, $p = .03$ (1.64% reduction in item variance). As expected, the variable's effect estimate was negative (see Figure 3.3)—the more stress friends a word had, the quicker participants were to respond. Similarly, comparing Models 5 and 7 showed that number of stress enemies contributed significantly to Model 7's fit, $\chi^2(1) = 7.12$, $p = .008$ (2.31% reduction in item variance). This effect of number of stress enemies was positive—the more enemies a word had, the slower participants were to respond. These findings indicate that, individually, number of stress friends and number of stress enemies predict lexical decision latencies. In the analyses that follow, we conduct a more stringent test of these effects by exploring the extent to which number of stress friends and number of stress enemies uniquely predict response latencies when the two variables are considered together.

First, we compared Model 6 (which includes number of stress friends but not number of stress enemies) with Model 8 (with includes both number of stress friends and stress enemies). This comparison tests whether a word's number of stress enemies contributes to the prediction of lexical decision latencies after accounting for its number of stress friends. Even with this additional control variable, number of stress enemies contributed significantly to Model 8's fit, $\chi^2(1) = 6.07$, $p = .01$ (1.98% reduction in item variance). As expected, this effect was positive; participants were slower to identify items with a large number of stress enemies, and were quicker to identify words with relatively few stress enemies.

Finally, we compared Model 7 (which includes number of stress enemies but not number of stress friends) with Model 8 (which includes both number of stress friends and stress enemies). The comparison showed that the fixed effect of number of stress friends contributed significantly to Model 8's fit, even after accounting for the effect of number of stress enemies (along with all other item-level and temporal controls). This was a significant, though marginal, effect ($\chi^2(1) = 3.86$, $p = .049$; 1.31% reduction in item variance). The direction of this effect was negative—words with many stress friends were recognized more quickly than those with few stress friends.

Taken together, these analyses show that adults make quicker lexical decisions when items have many stress friends, and at the same time make slower lexical decisions when items have many stress enemies. These effects are separable, as both number of stress friends and number of stress enemies survived inclusion of the other variable as a control. Figure 3.4 illustrates this, showing the marginal effects of number of stress friends and number of stress enemies as predictors in Model 8 with all other variables held constant at their means. As the figure shows, the effects of the two variables that were of numerically similar magnitude, but were in opposing directions. All told, these findings support the claim that readers draw on their exposure to all members of a stress neighbourhood when processing stress in multisyllabic written words.

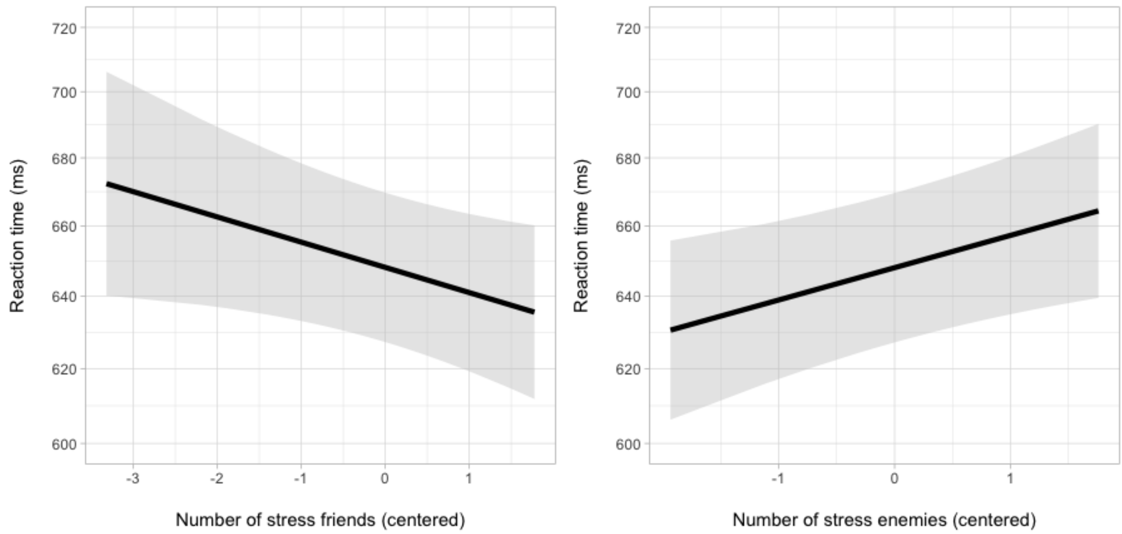


Figure 3.4. Marginal effects [95% CI] of a word's number of stress friends and number of stress enemies, respectively, as predictors of lexical decision latency in Model 8. All other predictors in Model 8 are set to their means. Graphs' *y*-axes are shown with log-transformed coordinates.

3.5 DISCUSSION

Growing behavioural evidence suggests that readers use word endings as cues to stress in written words (e.g., Arciuli & Cupples, 2006; Colombo & Sulpizio, 2015; Kelly et al., 1998; Mundy & Carroll, 2013; Sulpizio et al., 2013). Although statistical learning has been invoked as the mechanism by which readers become sensitive to the links between word endings and lexical stress in their language (e.g., Arciuli & von Koss Torkildsen, 2012; Colombo et al., 2014; Ševa et al., 2009), direct evidence for this claim has been lacking. To address this gap, the current study tested predictions that were derived from statistical learning principles, focusing on the claim that exposure to text input drives readers' sensitivity to the statistical patterns within it (per, e.g., Perruchet &

Pacton, 2006). English-speaking adults completed a visual lexical decision task with items that varied in their proportion of stress friends, number of stress friends, and number of stress enemies. Surprisingly, after accounting for proportion of stress friends, participants were no quicker to respond to words with many stress friends than those with fewer stress friends. This contradicts the claim that readers' sensitivity to stress neighbourhoods depends in part on how widely a given word ending occurs in the language (e.g., Sulpizio et al., 2013). However, proportion of stress friends was a robust predictor of readers' lexical decision performance, able to account for variance in response times after controlling for number of stress friends. Our second research question sheds light on the sources of this effect. Number of stress friends and number of stress enemies separately predicted adults' lexical decision latencies, with faster responses when words had many stress friends and slower responses when words had many stress enemies. Taken together, these results suggest that readers' use of word endings as cues to stress depends on their exposure to all words in a stress neighbourhood—both the positive examples of an association between the ending and stress pattern (stress friends) and the negative counterexamples of that association (stress enemies).

3.5.1 DIFFERENCES IN READERS' EXPOSURE TO STRESS NEIGHBOURHOODS

Our first set of analyses confirmed that proportion of stress friends—which captures the strength of a word ending's association with stress—plays a role in adults' visual word recognition. Specifically, we found that lexical decisions were made more quickly when words had a relatively large proportion of stress friends than when words had a smaller proportion of stress friends. This aligns nicely with the findings of prior studies (e.g., Kelly et al., 1998; Mundy & Carroll, 2013) and bolsters the existing work by showing an effect of proportion of stress friends after accounting for a word's number of stress friends—a control that was missing from previous examinations of stress neighbourhood consistency in English. Our study therefore affirms that, although proportion of stress friends cannot capture differences in exposure to a word ending's association with stress, it captures a salient feature of stress neighbourhoods that affects readers' processing of multisyllabic written words.

A more surprising finding was that number of stress friends did not predict lexical decision latencies after accounting for a word's proportion of stress friends. Our expectation to the contrary was based on the premise that readers are exposed to more examples of an ending's association with stress when that ending is reflected many stress friends (e.g., *NOVel*) than when it is reflected in relatively few (e.g., *ballOON*).[9] As such, finding that number of stress friends affects word recognition independently of proportion of stress friends would have provided much-needed direct evidence for the claim that statistical learning is a mechanism behind readers' sensitivity to endings as stress cues. Further support for our original prediction comes from the one existing study that speaks to this question (Sulpizio et al., 2013). In Sulpizio and colleagues' post hoc analysis of Italian pseudoword naming, both proportion and number of stress friends predicted adults' stress assignment; more notably, number of stress friends was a stronger predictor than proportion. The contrasting findings across these two studies leave us with

[9] Another viable way to capture readers' exposure to written stress cues is with a word's summed frequency of stress friends. In our data, the same pattern of results emerged when number of stress friends was replaced with summed frequency of stress friends. This reanalysis supports our overall conclusions, but should be taken with some caution as items were not initially selected with summed frequency of stress friends in mind.

questions regarding the extent to which readers' exposure to written stress cues drives their sensitivity to them.

Certainly, there are differences between Sulpizio and colleagues' (2013) study and our own that likely contribute to their discrepant findings. For instance, pseudoword naming and visual lexical decision tasks impose different task demands (e.g., phonological decoding; differentiating words from pseudowords) and involve different behavioural outcomes (binary stress placement accuracy; lexical decision latency). Differences between Italian and English themselves might also affect how readers use word endings as stress cues in their respective languages. While the nature of any such language difference is unclear empirically, the possibility that language differences influence readers' use of written stress cues would fit with conflicting findings in the literature to date. Specifically, studies in English have reliably found effects of stress neighbourhood consistency on lexical decision performance (Kelly et al., 1998; Mundy & Carroll, 2013), whereas studies in Italian generally find these consistency effects in naming tasks, but not lexical decision (e.g., Burani & Arduino, 2004; Colombo & Sulpizio, 2015). As a result of these clear differences, further research is warranted to reconcile our study's findings with those of Sulpizio and colleagues (2013). Even so, our results speak against a broad claim that, to guide stress placement, a word ending "not only has to be associated with stress in a relatively high proportion of cases but also has to appear in a large number of stress neighbours" (Sulpizio et al., 2013, p. 61).

From a technical standpoint, we argue that the current study's approach is particularly well-suited to comparing the effects of a word's proportion and number of stress friends on readers' processing of multisyllabic words. Our use of real word items allowed for our proportion of stress friends variable to range from very low (5%) to very high (100%) values, whereas Sulpizio and colleagues' pseudoword task necessarily limited them to high proportions (ranging from 65% to 100%). This restricted range may have weakened the ability of their proportion of stress friends variable to predict stress assignment, leaving more room for number of stress friends to do so. In contrast, we found proportion of stress friends to be a robust predictor of lexical decision latency when using a wider range of scores.

While we may have a technical explanation to account for the differing results across studies, we are still left with conceptual questions as to why participants in our study were no faster to respond to words with many stress friends than to those with few. When interpreting this finding, it is worth noting that statistical learning occurs over a protracted period; sensitivity to the regularities in written language builds gradually as readers encounter them repeatedly in text (e.g., Arciuli & von Koss Torkildsen, 2012). Our study focused on adults, whose performance reflects the knowledge that they have accumulated through years of reading text. Throughout those years, adults' cumulative exposure to text may have brought them to a point where their sensitivity to the characteristics of all stress neighbourhoods—including relatively small ones—is stable. This would help to explain why, once proportion of stress friends had been accounted for, our adult participants were equally quick to recognize words with rarely-occurring and commonly-occurring endings. While speculative, this possibility is consistent with evidence for diminishing returns in lexical characteristics' behavioural effects. For example, words that are relatively frequent tend to be recognized more quickly than less frequent words, but this effect diminishes to the point where it is barely apparent at high-frequency ranges (e.g., Adelman, 2012). If our speculation is correct, we would expect differences in stress neighbourhood size to affect word recognition among developing readers, who are still in the process of discovering the associations between word endings and lexical stress in their language (e.g., Arciuli et al., 2010). Developmental work that tests this prediction is thus a promising avenue for further research.

3.5.2 THE COMPONENTS OF STRESS NEIGHBOURHOODS

Our second research question explored the components of stress neighbourhoods that affect readers' processing of disyllabic written words, disentangling the respective effects of exposure to stress friends and exposure to stress enemies. Our results straightforwardly show that readers draw on all members of a stress neighbourhood during visual word recognition. After accounting for number of stress enemies, participants responded more quickly to words with many stress friends than those with few. At the same time, after accounting for number of stress friends, participants were slower to respond to words with many stress enemies than those with few. These findings offer the first empirical look at the specific regularities that are relevant to adults' processing of stress neighbourhoods in English. As such, they clarify the nature of adults' sensitivity to the associations between word endings and lexical stress more precisely than has been possible in previous studies (e.g., Kelly et al., 1998; Mundy & Carroll, 2013). Our results also validate interpretations that have been offered for the effect of stress neighbourhood consistency (as quantified with proportion of stress friends) on readers' word recognition. Specifically, we show that Kelly and colleagues (1998) were right to propose that stress friends facilitate word recognition while stress enemies impede it. Neither variable alone can fully account for the effects of written word endings as cues to stress.

These behavioural findings mirror the process by which connectionist models of word reading—which implement statistical learning as their mechanism—learn of the links between orthography and stress in disyllabic words (Arciuli et al., 2010; Perry et al., 2010; Ševa et al., 2009). Connectionist models are trained through exposure to a large number of words representative of the lexicon, but are given no direct guidance on the features of those words that are relevant to stress assignment. Among the presented words are some that serve as positive examples of a given regularity (e.g., *NOVel*, which illustrates the association between *–el* and first-syllable stress) and others that serve as counterexamples to that regularity (e.g., *hoTEL*). The models weigh their encounters with these words, yielding graded associations between spelling patterns and lexical stress that are subsequently used to place stress in novel words. Although connectionist architectures do not specify the regularities to which the model should attend, they are conceptually built on the notion that all exposures to written words can modify the weights between a given regularity (e.g., word endings' spellings) and outcome (e.g., stress placement). Our results suggest that this framework is a plausible account of how readers learn of the links between word endings and lexical stress in English.

3.5.3 THEORETICAL IMPLICATIONS

Our goal for this study was to explore the claim that statistical learning is a mechanism behind readers' sensitivity to associations between word endings and lexical stress (e.g., Arciuli & von Koss Torkildsen, 2012; Colombo et al., 2014). Although several of our findings do align with this theoretical framework, results were mixed. On the one hand, we argue that the separable effects of stress friends and stress enemies on word recognition aligns with predictions of the statistical learning framework. Both variables had graded effects on word recognition, and the effects were numerically of similar magnitude; together, this suggests that readers weigh their exposure to all members of a stress neighbourhood when learning of the associations between word endings and lexical stress in English. This is particularly important given that statistical learning is a general mechanism that cannot speak to the precise features of stress neighbourhoods that might inform readers' sensitivity to their regularities. As such, our findings help to clarify how readers' exposure to stress neighbourhoods affects their processing of multisyllabic written words. On the other hand, we did not find support for our strongest statistical learning prediction—that having many stress friends would facilitate word recognition after accounting for differences in words' proportion of stress friends. Instead, lexical decision latencies were similar for words from large and small stress neighbourhoods. This finding contradicts our expectation that readers would be most sensitive to the written stress cues that they encounter most often. Thus, further research is needed before we can draw firm conclusions about the role of statistical learning in readers' sensitivity to word endings as cues to stress in English.

3.5.4 FUTURE DIRECTIONS

The current study takes novel steps toward understanding the role of exposure to stress neighbourhoods in readers' sensitivity to word endings as stress cues. Building on this foundation, future work might consider other ways to quantify the extent of readers' exposure to the regularities in text. There are several viable ways to accomplish this goal. One possibility is to focus on elementary school-aged children rather than adults. As noted above, developing readers are in the process of becoming sensitive to the links between word endings and lexical stress (e.g., Arciuli et al., 2010), and thus can provide key insight into the potential effects of exposure to stress neighbourhoods early in the learning process. Another avenue for future work is to explore individual differences in readers' exposure to text alongside item-level differences in exposure to specific stress neighbourhoods. Just as some written word endings are encountered more frequently because they appear often in text, some individuals encounter more examples of those written word endings because they have greater overall levels of print exposure (e.g., Acheson, Wells, & MacDonald, 2008; Moore & Gordon, 2015; Stanovich & West, 1989). In the context of statistical learning, the extent to which statistical patterns in the language affect word processing likely interacts with an individual's level of exposure to text (Falkauskas & Kuperman, 2015). Finally, extensions of the current study that involve a larger and more varied set of items might allow for more fine-grained examinations of the claim that statistical learning drives readers' sensitivity to word endings' associations with stress. For example, a study whose items include a wider range of numbers of stress friends across the range of proportions would be well equipped to test for interactions between the two variables, further exploring the claim that readers are most sensitive to the characteristics of stress neighbourhoods that they encounter most often.

With these opportunities for future research in mind, the current study offers new insight into adult readers' sensitivity to word endings as cues to stress in English, exploring the role of statistical learning as a mechanism for that sensitivity. Our results confirm that word endings' associations with stress affect the speed of adults' visual word recognition, confirming prior work (e.g., Kelly et al., 1998; Mundy & Carroll, 2013). This points to a robust role for proportion of stress friends in influencing skilled word recognition, which, interestingly, seems to apply across stress neighbourhoods of differing size. Our findings also clarify the sources of this effect, showing that it is driven by the competing effects of a word's number of stress friends and its number of stress enemies. This suggests readers draw on their exposure to all members of a stress neighbourhood during visual word recognition. Taken together, these findings add to our growing understanding of the role of lexical stress in processing multisyllabic written words.

3.6 APPENDIX D: LEXICAL DECISION TASK ITEMS

A list of the real words included in the lexical decision task, along with each item's mean (and standard deviation) reaction time and accuracy scores across participants.

Item	Reaction time (ms)	Accuracy	Item	Reaction time (ms)	Accuracy
absolve	706.7 (195.7)	0.73 (0.45)	*cartoon*	643.0 (248.2)	1.00 (0.00)
actor	636.1 (154.2)	0.98 (0.13)	*caution*	616.4 (125.2)	1.00 (0.00)
acute	702.6 (157.3)	0.96 (0.19)	*climate*	603.3 (104.4)	1.00 (0.00)
advance	620.5 (163.4)	1.00 (0.00)	*compel*	752.6 (203.5)	0.68 (0.47)
advice	626.0 (167.0)	1.00 (0.00)	*complain*	640.9 (164.2)	0.98 (0.13)
album	646.7 (174.7)	0.96 (0.19)	*complete*	625.7 (265.1)	0.96 (0.19)
amid	791.3 (201.2)	0.64 (0.48)	*conceive*	651.4 (143.6)	1.00 (0.00)
antique	686.7 (218.2)	0.96 (0.19)	*conclude*	638.6 (187.1)	0.96 (0.19)
arcade	694.5 (145.6)	0.86 (0.35)	*consume*	655.4 (174.6)	0.98 (0.13)
arrest	602.8 (130.6)	0.98 (0.13)	*contact*	597.7 (154.5)	1.00 (0.00)
athlete	616.4 (151.0)	0.96 (0.19)	*costume*	606.8 (142.8)	1.00 (0.00)
attract	647.1 (201.7)	1.00 (0.00)	*critique*	682.1 (238.7)	0.98 (0.13)
balloon	603.8 (127.9)	1.00 (0.00)	*culture*	603.7 (108.2)	1.00 (0.00)
barrel	712.1 (328.7)	0.95 (0.23)	*debate*	629.2 (181.9)	1.00 (0.00)
below	591.9 (114.9)	0.95 (0.23)	*decade*	642.9 (139.5)	1.00 (0.00)
bestow	784.5 (170.0)	0.64 (0.48)	*deceive*	667.8 (161.7)	0.96 (0.19)
boutique	689.0 (114.2)	0.86 (0.35)	*deplete*	819.0 (273.1)	0.80 (0.40)
brigade	774.7 (148.5)	0.71 (0.46)	*despite*	684.9 (209.9)	0.91 (0.29)
bullet	607.6 (123.9)	0.98 (0.13)	*detest*	736.5 (199.5)	0.96 (0.19)
cadet	753.9 (150.4)	0.84 (0.37)	*device*	646.9 (177.2)	0.98 (0.13)
canal	726.5 (221.1)	0.93 (0.26)	*devout*	826.6 (210.6)	0.55 (0.50)
carpet	639.8 (378.1)	0.98 (0.13)	*diet*	598.2 (101.4)	0.98 (0.13)

Item	Reaction time (ms)	Accuracy	Item	Reaction time (ms)	Accuracy
discrete	664.2 (162.2)	0.98 (0.13)	*glucose*	709.8 (290.6)	0.91 (0.29)
dispose	680.2 (184.5)	1.00 (0.00)	*granite*	740.1 (181.9)	0.88 (0.33)
dispute	659.3 (136.6)	0.91 (0.29)	*grenade*	714.6 (182.7)	0.95 (0.23)
dissolve	634.4 (122.4)	0.84 (0.37)	*guidance*	659.3 (149.4)	0.96 (0.19)
donor	691.5 (165.0)	0.96 (0.19)	*harpoon*	811.7 (376.0)	0.70 (0.46)
dugout	877.6 (253.5)	0.86 (0.35)	*harvest*	612.6 (116.6)	0.98 (0.13)
enclose	682.4 (188.4)	0.93 (0.26)	*heinous*	792.5 (165.0)	0.61 (0.49)
endow	854.1 (286.7)	0.66 (0.48)	*hollow*	670.7 (150.7)	0.98 (0.13)
enhance	637.2 (116.5)	0.98 (0.13)	*honest*	588.1 (138.0)	1.00 (0.00)
ensue	754.1 (177.1)	0.84 (0.37)	*horror*	571.9 (139.2)	0.96 (0.19)
error	618.9 (148.8)	0.96 (0.19)	*intact*	673.2 (210.8)	1.00 (0.00)
exact	672.6 (313.7)	1.00 (0.00)	*intrude*	687.6 (179.2)	0.93 (0.26)
exclude	675.4 (158.5)	1.00 (0.00)	*invade*	653.3 (156.6)	1.00 (0.00)
extent	696.0 (157.4)	0.96 (0.19)	*invent*	660.5 (137.9)	0.95 (0.23)
exude	813.9 (223.8)	0.59 (0.50)	*invest*	633.6 (174.5)	0.95 (0.23)
fallout	640.6 (126.1)	0.95 (0.23)	*jacket*	645.9 (146.2)	0.96 (0.19)
famous	597.9 (116.2)	0.96 (0.19)	*jealous*	612.2 (142.5)	0.98 (0.13)
finite	735.8 (129.5)	0.84 (0.37)	*joyous*	719.1 (162.5)	0.93 (0.26)
fluid	649.1 (163.1)	1.00 (0.00)	*justice*	605.0 (143.4)	0.98 (0.13)
forbid	708.2 (215.7)	0.91 (0.29)	*label*	647.2 (153.7)	0.98 (0.13)
forum	739.8 (149.0)	0.88 (0.33)	*lagoon*	757.0 (180.6)	0.88 (0.33)
fragrance	646.6 (136.3)	0.98 (0.13)	*lament*	791.9 (193.8)	0.66 (0.48)
function	616.1 (131.2)	0.98 (0.13)	*lecture*	620.2 (125.7)	1.00 (0.00)
garlic	641.2 (165.5)	1.00 (0.00)	*legal*	588.1 (114.9)	0.98 (0.13)
garment	724.0 (192.3)	0.86 (0.35)	*logic*	582.4 (127.3)	1.00 (0.00)

Item	Reaction time (ms)	Accuracy	Item	Reaction time (ms)	Accuracy
lookout	638.9 (197.5)	0.98 (0.13)	*polite*	633.4 (179.0)	1.00 (0.00)
magic	579.5 (211.3)	1.00 (0.00)	*pollute*	690.6 (166.8)	0.88 (0.33)
malice	744.2 (145.7)	0.77 (0.43)	*prelude*	738.5 (221.2)	0.96 (0.19)
maroon	715.6 (249.0)	0.93 (0.26)	*presume*	703.5 (191.9)	0.96 (0.19)
mature	627.4 (143.5)	0.98 (0.13)	*propose*	674.5 (390.6)	1.00 (0.00)
mirror	580.1 (145.5)	1.00 (0.00)	*protrude*	811.4 (206.5)	0.80 (0.40)
mission	603.9 (141.0)	1.00 (0.00)	*prudent*	756.8 (167.6)	0.89 (0.31)
modest	626.1 (139.8)	0.95 (0.23)	*pursue*	643.2 (133.2)	1.00 (0.00)
mountain	633.9 (228.6)	1.00 (0.00)	*rapid*	613.6 (163.4)	0.96 (0.19)
mystique	714.7 (202.7)	0.89 (0.31)	*react*	643.9 (185.1)	0.98 (0.13)
nervous	600.0 (127.6)	0.98 (0.13)	*region*	628.7 (123.2)	0.96 (0.19)
neutral	668.9 (221.8)	0.98 (0.13)	*regret*	643.1 (155.2)	0.98 (0.13)
notion	705.4 (196.0)	0.91 (0.29)	*relate*	641.2 (168.4)	0.98 (0.13)
novel	632.0 (184.1)	0.96 (0.19)	*rescue*	622.6 (120.3)	0.96 (0.19)
novice	705.6 (161.7)	0.89 (0.31)	*retain*	712.6 (219.5)	0.93 (0.26)
nuisance	794.6 (204.3)	0.84 (0.37)	*revolve*	735.8 (226.3)	0.86 (0.35)
obtain	624.9 (156.9)	1.00 (0.00)	*royal*	629.3 (163.2)	1.00 (0.00)
onion	639.9 (176.0)	0.96 (0.19)	*saloon*	685.7 (179.1)	0.84 (0.37)
oppose	662.5 (132.1)	0.98 (0.13)	*salute*	693.2 (291.3)	0.93 (0.26)
ornate	809.3 (213.3)	0.68 (0.47)	*secrete*	679.7 (169.7)	0.91 (0.29)
oval	669.6 (133.8)	0.95 (0.23)	*secure*	618.3 (145.1)	0.98 (0.13)
parade	676.1 (310.5)	0.98 (0.13)	*seizure*	658.3 (117.0)	0.95 (0.23)
perceive	666.6 (147.4)	1.00 (0.00)	*senate*	761.1 (199.2)	0.96 (0.19)
pirate	666.6 (182.2)	1.00 (0.00)	*solid*	571.4 (101.9)	1.00 (0.00)
plastic	600.8 (96.0)	1.00 (0.00)	*spectrum*	664.1 (127.4)	0.96 (0.19)

Item	Reaction time (ms)	Accuracy	Item	Reaction time (ms)	Accuracy
statute	669.3 (187.2)	0.82 (0.39)	*tribute*	669.0 (171.2)	1.00 (0.00)
substance	627.3 (174.5)	0.98 (0.13)	*unique*	611.0 (124.7)	0.98 (0.13)
subtract	701.7 (307.0)	1.00 (0.00)	*unite*	685.4 (302.6)	0.96 (0.19)
suffice	695.1 (145.2)	0.95 (0.23)	*valid*	653.8 (268.3)	0.98 (0.13)
sustain	678.3 (162.9)	0.95 (0.23)	*venture*	750.0 (333.1)	0.93 (0.26)
swallow	618.4 (139.4)	1.00 (0.00)	*venue*	679.4 (200.6)	0.89 (0.31)
tantrum	763.6 (150.5)	0.93 (0.26)	*vessel*	681.2 (132.2)	0.93 (0.26)
technique	620.1 (142.6)	1.00 (0.00)	*villain*	715.7 (231.8)	0.95 (0.23)
throughout	689.9 (202.2)	0.96 (0.19)	*virtue*	659.6 (141.8)	0.95 (0.23)
tissue	626.9 (143.0)	0.98 (0.13)	*volume*	585.4 (99.9)	0.96 (0.19)
topic	602.0 (152.6)	1.00 (0.00)	*widow*	656.2 (181.6)	0.98 (0.13)
traffic	592.7 (127.6)	0.98 (0.13)	*wondrous*	755.0 (186.5)	0.82 (0.39)
treatment	633.1 (145.2)	1.00 (0.00)	*workout*	641.1 (201.0)	0.98 (0.13)

3.7 APPENDIX E: MODELS' RANDOM EFFECTS STRUCTURES

Table 3.2.

Summary of the variance components (standard deviations) for each random-effects parameter in Models 1 to 4 (addressing research question #1).

	Model 1: Control model	Model 2: Proportion of stress friends	Model 3: Number of stress friends	Model 4: Full model
Random intercepts				
Item	0.028597	0.027569	0.028122	0.027560
Participant	0.051443	0.051463	0.051450	0.051463
Item random slopes				
Word frequency	0.004736	0.004730	0.004728	0.004729
Word length	0.003543	0.003518	0.003528	0.003517
OLD	0.004068	0.004066	0.004073	0.004067
Mean bigram frequency	—	—	—	—
Proportion stress friends	—	—	—	—
Number stress friends	—	—	—	—
Participant random slopes				
Trial number	0.004040	0.004053	0.004028	0.004050
Preceding RT	0.006857	0.006865	0.006870	0.006866
Correlation parameters	—	—	—	—
Residuals	0.078655	0.078654	0.078655	0.078654

Notes. Empty cells indicate random effects that were excluded during the model selection process (Bates et al., 2015). OLD = Orthographic Levenshtein Distance. Preceding RT = participant's response time to the preceding trial. Correlation parameters = correlations between random intercepts and slopes.

Table 3.3.

Summary of the variance components (standard deviations) for each random-effects parameter in Models 5 to 8 (addressing research question #2).

	Model 5: Control model	Model 6: Number of stress friends	Model 7: Number of stress enemies	Model 8: Full model
Random intercepts				
Item	0.028575	0.028105	0.027916	0.027549
Participant	0.051465	0.051472	0.051475	0.051481
Item random slopes				
Word frequency	0.004835	0.004827	0.004826	0.004820
Word length	0.003493	0.003479	0.003483	0.003471
OLD	0.003881	0.003888	0.003886	0.003891
Mean bigram frequency	—	—	—	—
Number of stress friends	—	—	—	—
Number stress enemies	0.003080	0.003061	0.002950	0.002949
Participant random slopes				
Trial number	0.004070	0.004057	0.004080	0.004067
Preceding RT	0.006846	0.006859	0.006846	0.006858
Correlation parameters	—	—	—	—
Residuals	0.078601	0.078601	0.078603	0.078603

Notes. Empty cells indicate random effects that were excluded during the model selection process (Bates et al., 2015). OLD = Orthographic Levenshtein Distance. Preceding RT = participant's response time to the preceding trial. Correlation parameters = correlations between random intercepts and slopes.

CHAPTER 4
STRESS GUIDES SPELLING DECISIONS INDEPENDENTLY OF VOWEL QUALITY: UNTANGLING THE GORDIAN KNOT

4.1 ABSTRACT

Lexical stress is tightly connected with vowel quality in English—a link that has hindered efforts to establish a specific role for stress in processing written English words. In the current study, we addressed this challenge with a task that manipulates stress independently of vowel quality, disentangling the two factors to determine whether adults and school-aged children use lexical stress to guide their pseudoword spelling decisions. Participants completed a spelling choice task that offered multiple spellings for an aurally-presented disyllabic pseudoword. Crucially, one spelling's ending was statistically associated with first-syllable stress (e.g., *–et*), and another was associated with second-syllable stress (e.g., *–ette*). Adults preferentially chose the spelling associated with the stress pattern of each aurally-presented pseudoword, as did children in the late-elementary grades (Grades 5–6). By contrast, younger children (Grades 1–4) preferred the more common ending spellings (e.g., *–et*), regardless of stress pattern. Empirically, these findings further our understanding of readers' sensitivity to the links between stress and spelling in English. In terms of theory, they inform competing models of the process by which individuals choose among spelling alternatives.

4.2 INTRODUCTION

A challenge of the English writing system arises from the fact that many individual sounds and sound sequences can be spelled in multiple ways. The ending sequence /ɛt/, for example, can plausibly be spelled *–et*, as in the word *asset*, or spelled

–ette, as in *roulette*. Similarly, the ending /i/ is spelled differently in *trusty* and *trustee*, an instance in which the spelling distinction is particularly important given that it marks a lexical contrast. This type of sound-to-spelling inconsistency is common in English, both at the level of individual phonemes and when mapping sound to spelling in larger sound units (e.g., Kessler & Treiman, 2001; Ziegler, Stone, & Jacobs, 1997). As a result, English spellers are frequently faced with making a choice between different spelling alternatives. Existing models of spelling give contrasting accounts of how these spelling choices are resolved (Barry & Seymour, 1988; Houghton & Zorzi, 2003). Yet because these models focus largely on monosyllabic words, they overlook lexical stress—a key feature of multisyllabic English words, which make up over 90% of the lexicon (Baayen, Piepenbrock, & van Rijn, 1995). In the current study, we explore the possibility that lexical stress may help to guide English speakers' choice of spelling for multisyllabic pseudowords. We look at whether stress informs the spelling choices made by skilled adult spellers, and also explore whether this is the case for developing spellers. In doing so, we test the claim that readers of English are sensitive to the statistical associations between spelling and lexical stress, making use of those associations when they process written words.

Although the English writing system does not mark stress placement directly, recent evidence suggests that it contains probabilistic cues to stress. Corpus analyses of the texts read by children and adults have established that spelling sequences—ending spellings in particular—are often associated with specific stress patterns (e.g., Arciuli & Cupples, 2006; Arciuli, Monaghan, & Ševa, 2010; Kelly, Morris, & Verrekia, 1998; Monaghan, Arciuli, & Ševa, 2016). For example, the ending *–et* is strongly associated with first-syllable stress, as that is the stress pattern of approximately 90% of the disyllabic English words that end with that spelling sequence. Similarly, approximately 94% of words ending in *–oon* have second-syllable stress. While these examples are quite strongly associated with their respective stress patterns, the strength of associations between spelling and stress varies along a continuum across endings (Arciuli & Cupples, 2006). Yet when strong associations exist, a word's spelling can serve as a cue to its stress pattern, and in turn, lexical stress may serve as a cue to spelling.

4.2.1 READERS' USE OF THE ASSOCIATIONS BETWEEN ENGLISH STRESS AND SPELLING

A growing body of behavioural research suggests that readers of English use the associations between ending spellings and stress during visual word recognition. For example, in naming and lexical decision tasks, adults respond more quickly to words with ending spellings that are strongly associated with their stress patterns (e.g., *COMet*, *rouLETTE*) than ones with endings that are not (e.g., *caDET*, *PALette*; Kelly et al., 1998; Mundy & Carroll, 2013). This has been taken to suggest that ending spellings influence readers' processing of familiar words. Similarly, ending spellings seem to guide stress placement in unfamiliar pseudowords. Arciuli and Cupples (2006) demonstrated this by presenting English-speaking adults with disyllabic written pseudowords whose endings were associated with either first- or second-syllable stress, asking them to indicate which syllable should receive emphasis. Participants tended to assign the stress pattern associated with each pseudoword's ending spelling (see e.g., Sulpizio, Arduino, Paizi, & Burani, 2013, for similar studies in Italian). In other words, adult readers seem to be sensitive to the associations between stress and spelling in English, and they make use of those associations when processing written words.

Relatively little research has explored whether this is also true of developing readers. In one existing study of English-speaking children, 5- to 12-year-olds were shown written pseudowords whose ending spellings are associated with either first- or second-syllable stress in children's texts (Arciuli et al., 2010). The children were asked to read each pseudoword aloud, and their disyllabic pronunciations were coded for stress placement. The youngest children (ages 5 to 8 years) made slight use of the words' endings to guide stress placement, but pronounced most pseudowords with first-syllable stress, regardless of ending. This tendency reflects the overall predominance of words with first-syllable stress in English, and suggests that young children draw on the overall distribution of stress patterns in the language without as much regard to the more fine-grained links between spelling and stress. On the other hand, children aged 9 years and older made use of spelling to guide stress placement, and tended to pronounce the pseudowords in accordance with their ending spellings. A similar developmental pattern has been reported in Italian (see Colombo, Deguchi, & Boureux, 2014; Sulpizio & Colombo, 2013). By the mid-elementary school years, then, children seem to be sensitive to the association between written word endings and lexical stress in English.

While the work described above is compelling, the methods employed to date cannot fully disentangle the effects of stress from those of vowel quality. In English, lexical stress and vowel quality are highly related; unstressed syllables often include vowels that are reduced to the phoneme schwa (/ə/; Halle & Vergnaud, 1987). The correlation between stress and vowels in English makes their respective effects hard to distinguish, presenting a challenge for many widely-used experimental paradigms. For instance, studies to date using real English words (e.g., Kelly et al., 1998) allow the effects of stress and vowel quality to co-occur, as they do in the language. Consequently, any stress errors that participants make also tend to involve errors in vowel quality. For example, when the word *canal* is correctly pronounced with its second-syllable stress pattern, its phonemic pronunciation is /kə 'NÆL/. With stress misplaced to the first syllable (which is the pattern suggested by the ending *–al*), the vowels change to /'KÆN əl/. As Kelly and colleagues (1998) acknowledge, these vowel mismatches could, in principle, be what drives the delayed lexical decisions that they attribute to lexical stress. This view would not dispute the presence of associations between spelling and stress, but might undercut the apparent role that stress plays in word recognition.

A similar issue arises in studies that use pseudoword naming, where stress and vowel quality are generated by the participant and thus are outside of the researchers' control. In Arciuli and colleagues' (2010) study with children, for instance, participants were asked to pronounce pseudowords like *bedoon*. By mid-childhood, they tended to produce that item with second-syllable stress. Certainly, this may indicate that children recognized the link between *–oon* and second-syllable stress. However, it could also be that they recognized *–oon* as an unlikely spelling for the ending of the word when pronounced with first-syllable stress as /'BƐD ən/. Taken together, then, the empirical research to date has not yet disentangled the effects of stress and vowel quality on readers' processing of multisyllabic written words. We see this issue as particularly noteworthy in light of work done on oral language, where prominent researchers have argued that vowel quality, and not lexical stress per se, is critical to recognizing words in English (see, e.g., Bond & Small, 1983; Cutler, 1986).

To address this confound in the current study, we explored whether stress patterns guide spelling decisions using a paradigm that controls for vowel quality. In this paradigm, we made use of a subset of word endings for which there are two possible spellings; critically, these spellings are phonemically similar to each other, but are associated with different stress patterns. The spellings *–et* and *–ette* are an example. While most English disyllables ending in *–et* have first-syllable stress (e.g., *ASSet*), most ending in *–ette* have second-syllable stress (e.g., *rouLETTE*)—the different spellings capture differences in lexical stress. In an experimental spelling choice task, these ending pairs can be presented to participants as possible spellings for an aurally-presented pseudoword. Importantly, this approach lets us manipulate stress patterns of stimuli without reducing unstressed vowels, thereby removing the potential influence of stress' key phonemic correlate. This allows for a stringent test of the claim that readers are sensitive to associations between ending spellings and lexical stress.

To our knowledge, this spelling choice paradigm has only been used once, in an unpublished study by Verrekia (1996). In that study, adult readers of English were presented with a series of spoken disyllabic pseudowords, pronounced with either first-syllable stress (/'KLÆG rɛt/) or second-syllable stress (/klæg 'RƐT/). Participants were then asked to select one of two forced-choice spelling alternatives that differed only in how their endings were spelled (*klagret*; *klagrette*). As expected, the aural stress manipulation informed participants' spelling choices. That is, adults were more inclined to choose the ending spelling associated with second-syllable stress after hearing pseudowords pronounced with second-syllable stress than ones pronounced with first-syllable stress. Verrekia's (1996) study suggests that adults do in fact use lexical stress to guide their spelling decisions, drawing on the statistical regularities that link stress and spelling when choosing among plausible alternatives for multisyllabic words.

In the current study, we adopted the same spelling choice paradigm. Our aim was to replicate Verrekia's key findings with adults and to extend them to developing readers. In doing so, our study contributes to the growing literature on readers' use of stress when processing written words.

4.2.2 THEORETICAL CONTRIBUTIONS OF UNDERSTANDING THE IMPACT OF LEXICAL STRESS ON SPELLING CHOICE

The spelling choice task that we used in the current study lets us touch on an issue at the heart of competing models of the spelling process (Barry & Seymour, 1988; Houghton & Zorzi, 2003). Existing models recognize that, although multiple spellings are often possible for a given sound sequence in English, those spellings are seldom equally probable. Critically, the models disagree on the factors that best establish the probability of readers' choices between different spelling alternatives. At a broad level, for instance, some spellings occur more frequently than others in English words (e.g., *–et* is more common overall than *–ette*). Some accounts view this as central to how spellers choose among competing alternatives, arguing that spellers will opt to use higher-frequency spellings for a given sound more often than lower-frequency ones (e.g., Barry & Seymour, 1988). In contrast, other researchers have noted that this approach ignores the fact that fine-grained contextual factors can inform spelling decisions (e.g., Treiman & Kessler, 2006). For example, both children and adults draw on characteristics of the preceding vowel when choosing among spellings (Hayes, Treiman, & Kessler, 2006). Importantly, this work shows that sensitivity to subtle language regularities can lead spellers to preferentially choose the option that is less common in the language overall.

Recent models of spelling have begun to incorporate some of these statistical regularities, allowing for characteristics of the word in which a sound appears to influence spelling decisions (Houghton & Zorzi, 2003).

Notably, the spelling models to date have not taken lexical stress into account. As a result, they are not equipped to fully capture regularities that apply to the spelling of multisyllabic words in English. By focusing on lexical stress, the current study fills a key gap in existing work on how spellers choose from among plausible spelling alternatives, and this may help to inform the development of more robust spelling models. For example, when faced with a word that ends with the sound sequence /ɛt/, we might find that spellers indiscriminately prefer the more common spelling *–et*. This would be consistent with views that emphasize the overall frequency of sound-to-spelling correspondences (Barry & Seymour, 1988). On the other hand, if spellers instead draw on more fine-grained statistical regularities in the writing system (Houghton & Zorzi, 2003), we might expect a word's lexical stress pattern to guide their spelling decisions. In that case, they would be more inclined to choose the spelling *–ette* for words that are pronounced with second-syllable stress than those pronounced with first-syllable stress (and vice versa for the spelling *–et*). Indeed, we might find support for both accounts at different points in development. By exploring the role of stress in spelling among both skilled readers and developing ones, the current study offers a crucial test of these possibilities.

4.2.3 THE CURRENT STUDY

The current study consists of two experiments that explore whether stress guides spelling decisions in adults (Experiment 1) and in elementary school-aged children

(Experiment 2). The benefits to looking at the effects of stress on spelling decisions are twofold. On an empirical level, this study builds on existing work into readers' sensitivity to the associations between word endings and lexical stress (e.g., Arciuli et al., 2010; Kelly et al., 1998). Critically, it does so with an experimental design that controls for vowel quality as a confounding factor. We use this controlled spelling choice paradigm with both adults and children to get a view of the factors that affect processing across reading development. In terms of theory, our work pits the overall frequency of spellings against more fine-grained spelling regularities, thereby testing predictions of competing spelling models (Barry & Seymour, 1988; Houghton & Zorzi, 2003).

Throughout this study, we refer to endings associated with first-syllable stress as simple spellings (e.g., *–et*), and those associated with second-syllable stress as extended spellings (e.g., *–ette*). This terminology is based on the observation that, broadly speaking, spellings associated with second-syllable stress tend to be longer than those associated with first-syllable stress, and use more letters than strictly necessary to represent the ending's phonemes. In contrast, endings associated with first-syllable stress tend to be simpler spellings that do not contain additional silent letters (Kelly et al., 1998; Verrekia, 1996). Thus, for clarity, we will use the terms simple and extended spellings to refer to written word endings, and will use the terms first- and second-syllable stress only when describing aurally-presented words.

Experiment 1 is based closely on Verrekia's (1996) study, and as such, we expect our results to be similar. That is, we expect that the aural stress patterns of pseudowords will affect adults' spelling choices. Specifically, we expect adults to preferentially choose extended spellings (e.g., *–ette*) after hearing pseudowords with second-syllable stress, and to preferentially choose simple spellings (e.g., *–et*) after hearing pronunciations with first-syllable stress. However, it also is possible that we might see a general preference for simple spellings (e.g., *–et*), regardless of stress pattern. This finding would be consistent with claims that spellers tend to prefer more frequent alternatives over less frequent ones (Barry & Seymour, 1988).

Experiment 2 uses the same spelling choice paradigm with children in Grades 1 to 6, spanning the elementary years. To get a broad sense of stress' effects on spelling across early reading development, we tested children in the early- (Grades 1 and 2), mid- (Grades 3 and 4), and late-elementary years (Grades 5 and 6). It seems unlikely that the youngest, early-elementary aged children will use stress to inform their spelling choices. Instead, we expect that children in Grades 1 and 2 will prefer the more-frequent simple spellings over the less-frequent extended ones, regardless of aural stress pattern (consistent with e.g., Barry & Seymour, 1988). We base this prediction on two key factors. First, as relatively new readers, children in the early elementary grades will have had less exposure to text than older children. Consequently, they will have had less opportunity to learn the various subtle regularities that govern English spelling. We therefore expect them to draw more on the broad frequencies of spellings than on the more fine-grained associations between stress and spelling. Moreover, in prior research, six- and seven-year-old children have shown limited sensitivity to the links between word endings and lexical stress in English (Arciuli et al., 2010).

In contrast, we do expect stress to inform spelling decisions among children in the mid-elementary and late-elementary grades. With their additional years of exposure to text, these children will have had more opportunity to associate words' spellings and stress patterns, making it plausible that these associations might override differences in spellings' broad frequencies (consistent with e.g., Houghton & Zorzi, 2003; Treiman & Kessler, 2006). This prediction is in line with the pattern we anticipate in adults, though it would be reasonable to expect the effects to be smaller among developing than skilled readers. Supporting this, the existing research suggests that children are sensitive to the links between word endings and stress at around age 9 (Arciuli et al., 2010). With that being said, it is possible that those prior findings were driven by vowel quality, rather than by lexical stress. If that is the case, we might not see any evidence that children use stress to guide their spelling decisions.

Taken together, the current study will allow us to (1) confirm that English-speaking adults use stress to inform their choice of spelling when presented with two phonemically plausible alternatives, and (2) establish if, and when, school-aged children begin to do the same.

4.3 EXPERIMENT 1: ADULT READERS

4.3.1 METHOD

Participants. Participants were 37 adult university students in eastern Canada (mean age = 20.3 years, *SD* = 1.69), most of whom were women (four participants were men; one did not specify a gender). One additional participant was tested, but was excluded from analyses as English was not their first language. As compensation, participants earned either $5 or 0.5 percentage points toward a university course.

Materials. The experimental spelling choice task used in this experiment involved items adapted from Verrekia (1996). Participants listened to a set of aurally-presented pseudowords; for each, they were asked to select their preferred spelling from among four forced-choice options.

Pseudoword items. Each aurally-presented pseudoword contained one of 14 word endings.[10] These endings are shown in Appendix F. The endings were drawn from available corpus analyses (Arciuli & Cupples, 2006; Kelly et al., 1998; Verrekia, 1996), and were selected according to two key criteria. First, each ending's pronunciation (e.g., /ɛt/) had at least two plausible spellings (e.g., *–et* and *–ette*). Second, the two spelling alternatives differed in their association with lexical stress, such that one spelling could be designated as indicating first-syllable stress (e.g., *–et*) and the other second-syllable stress (e.g., *–ette*). We refer to these written ending alternatives as simple spellings (*–et*; associated with first-syllable stress) and extended spellings (*–ette*; associated with second-syllable stress), respectively.

Characteristics of the 14 word endings are shown in Appendix F. Although the simple and extended spellings were chosen for their association with different stress patterns, the strength of that association was similar across conditions. That is, the proportion of words with our simple spellings that have first-syllable stress (M = .84) was equivalent to the proportion of words with our extended spellings that have second-syllable stress (M = .78; Mann-Whitney U = 76, p = .31). Consistent with observations made by Kelly and colleagues (1998), extended spellings (associated with second-syllable stress) contained significantly more letters than did simple spellings ($t(26)$ = 4.69, p < .001). Extended spellings were also less frequent, appearing in fewer mono- and

[10] Two additional two endings (/ɛl/ and /æs/) were presented to participants, but were excluded from analysis. An examination of the CELEX English database (Baayen et al., 1995) confirmed that both spellings used for those endings are predominantly associated with first-syllable stress, and they therefore do not satisfy our second selection criterion.

disyllabic English words than simple spellings (Mann-Whitney $U = 37.5$, $p = .005$).[11]

We created eight pseudowords for each ending. Following on Verrekia (1996), pseudowords were formed by generating an initial syllable that was unique to each item, plus a second-syllable initial consonant to which the ending was appended (e.g., /klæg/ + /r/ + /ɛt/). Items conformed to the phonotactic and orthographic constraints of English. A native English speaker recorded two audio versions of each item: one with first-syllable stress, and another with second-syllable stress (Audacity 2.0.3, Mazzoni & Dannenberg, 2013; 44.1 kHz sampling rate, 16 bit resolution, mono). Vowels in unstressed syllables were not reduced to schwa (i.e., /'DƐL kɑr/, not /'DƐL kər/), as doing so would have introduced vowel quality as a confounding factor (and rendered several of our ending spellings implausible; e.g., *–arre* for /ər/). Importantly, although vowel quality is a corollary of stress in English, Verrekia (1996) found that native listeners' stress perception was accurate in pseudowords without vowel reduction.

After recording the pseudoword stimuli, we conducted two sets of acoustic analyses to verify the key aspects of our manipulation. In the first set of analyses, we established each pseudoword's stress pattern by calculating normalized Pairwise Variability Indices, or PVIs (Ling, Grabe, & Nolan, 2000). The PVI is a flexible metric that can be used to capture differences in the acoustic features of two adjacent phonological units—including, notably, the features that differentiate stressed from unstressed syllables (Ballard, Djaja, Arciuli, James, & van Doorn, 2012). In English, stressed syllables tend to have vowels with a higher pitch, greater intensity, and longer

[11] In keeping with Barry and Seymour (1988), we have reported endings' frequencies across both monosyllabic and disyllabic words. However, the same frequency effect emerges when comparing simple and extended spelling frequencies in disyllabic words alone (Mann-Whitney $U = 37.0$, $p = .005$).

duration than those in unstressed syllables (e.g., Fry, 1958; note that these characteristics are separate from the vowel reduction described earlier). Normalized PVIs can be calculated separately for each of these acoustic features by comparing a measurement from the first syllable's vowel (e.g., its pitch in Hertz) with that from the second syllable's vowel. Specifically, the difference between the first and second vowels' acoustic measurements is divided by their average. As a result, the PVI can tell us the direction of a lexical stress contrast: positive scores indicate a word with first-syllable stress, whereas negative scores indicate second-syllable stress. PVI scores also quantify the magnitude of a stress contrast: higher absolute values indicate a larger degree of contrast between the stressed and unstressed syllable; scores at or near zero indicate no contrast.

For the current study, we calculated two PVI scores for each item, with one comparing the vowels' maximum pitch (Hz) and the other comparing their maximum intensity (dB). However, we did not compare vowel durations (cf. Ballard et al., 2012), because our stimuli were disyllabic and their second vowels therefore occurred in each item's final syllable. English vowels are significantly lengthened when they occur in word-final syllables, whether that syllable is stressed or unstressed (Oller, 1973). As such, we would expect to find negative PVI scores for vowel duration regardless of a pseudoword's stress pattern. For our purposes, then, vowel duration PVI scores did not provide a useful measure of stress contrastivity.

All acoustic measurements of pitch and intensity were made using Praat (Boersma & Weenink, 2017, version 6.0.25). Table 4.1 shows mean PVI scores, calculated separately for items with first- and second-syllable stress. One-sample *t*-tests showed that items pronounced with first-syllable stress had PVI scores that were significantly greater than zero for both pitch, $t(111) = 24.88$, $p < .001$, $d = 2.35$, and intensity, $t(111) = 15.39$, $p < .001$, $d = 1.45$. Similarly, items pronounced with second-syllable stress had PVI scores significantly lower than zero for both pitch, $t(111) = -16.71$, $p < .001$, $d = -1.58$, and intensity, $t(111) = -16.02$, $p < .001$, $d = -1.51$. These analyses help to confirm that our pseudowords were indeed produced with their intended stress patterns.

Table 4.1.

Mean (and SD*) PVI scores for pitch and intensity, quantifying the lexical stress contrasts in items pronounced with first-syllable stress and with second-syllable stress.*

	Experiment 1 (112 items)	Experiment 2 (48 items)
	Mean *(SD)*	Mean *(SD)*
PVI: Pitch		
First-syllable stress	44.08 (18.75)	39.08 (17.28)
Second-syllable stress	-22.29 (14.11)	-26.82 (18.38)
PVI: Intensity		
First-syllable stress	5.47 (3.76)	7.65 (3.46)
Second-syllable stress	-6.43 (4.25)	-6.81 (4.99)

Notes. PVI = normalized Pairwise Variability Index. PVI = $100 \times \{(a_1 - a_2) / [(a_1 + a_2) / 2]\}$, where a_1 and a_2 are the acoustic measurements being compared from the first and second vowels, respectively (Ling et al., 2000).

Our second acoustic analysis confirmed that vowel quality was controlled by comparing vowels across the two pronunciations of each pseudoword. For example, we compared the vowel /æ/ in /'KLÆG rɛt/ to the corresponding vowel in /klæg 'RƐT/, with the goal of ensuring that they were acoustically similar. To that end, we measured the first three formants of each vowel (F1, F2, and F3, respectively), as the combination of these formants effectively differentiates between vowels in English (Ladefoged &

Johnson, 2010). We calculated average formant frequencies (Hz) based on measurements taken at 10ms intervals across the full vowel duration, extracted through a semi-automatic script in Praat (Boersma & Weenink, 2017). Table 4.2 shows descriptive statistics for these formants across items pronounced with first- and second-syllable stress. To compare vowels, we conducted a repeated-measures MANOVA with the three formants (F1; F2; F3) entered as dependent variables and aural stress pattern (first-syllable stress; second-syllable stress) as a within-subjects factor. Doing so showed no effect of aural stress pattern on the combination of vowel formants, $\Lambda = .98$, $F(3, 218) = 1.44$, $p = .23$, $\eta_p^2 = .02$. This helps to confirm that vowel quality was similar across our pseudoword stimuli.

Table 4.2.

Mean (and SD*) formant frequencies of the vowels in pseudowords pronounced with first- and second-syllable stress.*

	Experiment 1 (112 items)	Experiment 2 (48 items)
	Mean *(SD)*	Mean *(SD)*
Vowels' F1 (Hz)		
First-syllable stress	674.81 (173.20)	656.61 (188.57)
Second-syllable stress	667.98 (179.59)	646.45 (181.98)
Vowels' F2 (Hz)		
First-syllable stress	1794.62 (436.65)	1858.40 (478.46)
Second-syllable stress	1789.30 (420.68)	1857.97 (459.25)
Vowels' F3 (Hz)		
First-syllable stress	2730.58 (279.21)	2761.68 (265.05)
Second-syllable stress	2743.48 (277.84)	2758.67 (289.09)

Notes. F1 = first formant; F2 = second formant; F3 = third formant.

Spelling choice task. To ensure that participants heard only one pronunciation of each item, the audio recordings were divided into two between-subjects lists, each with 128 total pseudowords. Within lists, half of the items for each ending were pronounced with first-syllable stress and half were pronounced with second-syllable stress. For example, approximately half of participants heard the items /'KLÆG rɛt/ and /vɛr 'DƐT/, whereas the other half heard /klæg 'RƐT/ and /'VƐR dɛt/.

For each aurally-presented item, participants were given four spellings to choose from. These options differed only in terms of how the endings were spelled: two were the simple and extended target spellings described above (e.g., *klagret*, *klagrette*) and two were distractor spellings (e.g., *klagrap*, *klagrappe*). Distractors maintained the target spellings' length and syllable structure, but were not plausible spellings of the pseudoword's pronunciation. Participants made their spelling choices in a printed booklet, with the four options for each item appearing on a single, numbered line. Items were divided into eight blocks of 16, and were distributed so that only one instance of each ending appeared per page. Items were randomized within blocks, but were presented in the same order across participants. Presentation order of the four spelling choices was counterbalanced across items.

Procedure. Participants were tested in groups of up to four. We ensured that participants could not see one another's responses when completing the task. Audio recordings were put into a continuous playlist with a five-second pause between items. Before testing began, the experimenter played a series of practice items to familiarize participants with the procedure and to ensure that the audio quality was adequate. Participants were then told that they would hear a series of non-words, and that for each, they had the option to choose one of four provided spellings. They were asked to circle the spelling that best represented what they heard. The five-second delay between items gave sufficient time for participants to respond, though they were told that the recording could be paused as needed. In total, the testing session lasted approximately 20 minutes.

4.3.2 RESULTS AND DISCUSSION

We calculated the rate at which participants chose the simple target spelling, the extended target spelling, or a distractor spelling for the presented pseudowords, with the two distractor options combined into one category. The right-hand panels of Figure 4.1 show these mean rates as a function of aural stress pattern. Unsurprisingly, participants generally limited their selections to the two target spellings; distractors were chosen on less than 0.2% of all trials. Given their low selection rate and the likelihood that distractor selections reflect cases where participants misheard the item, we excluded the distractor spellings from further analysis.

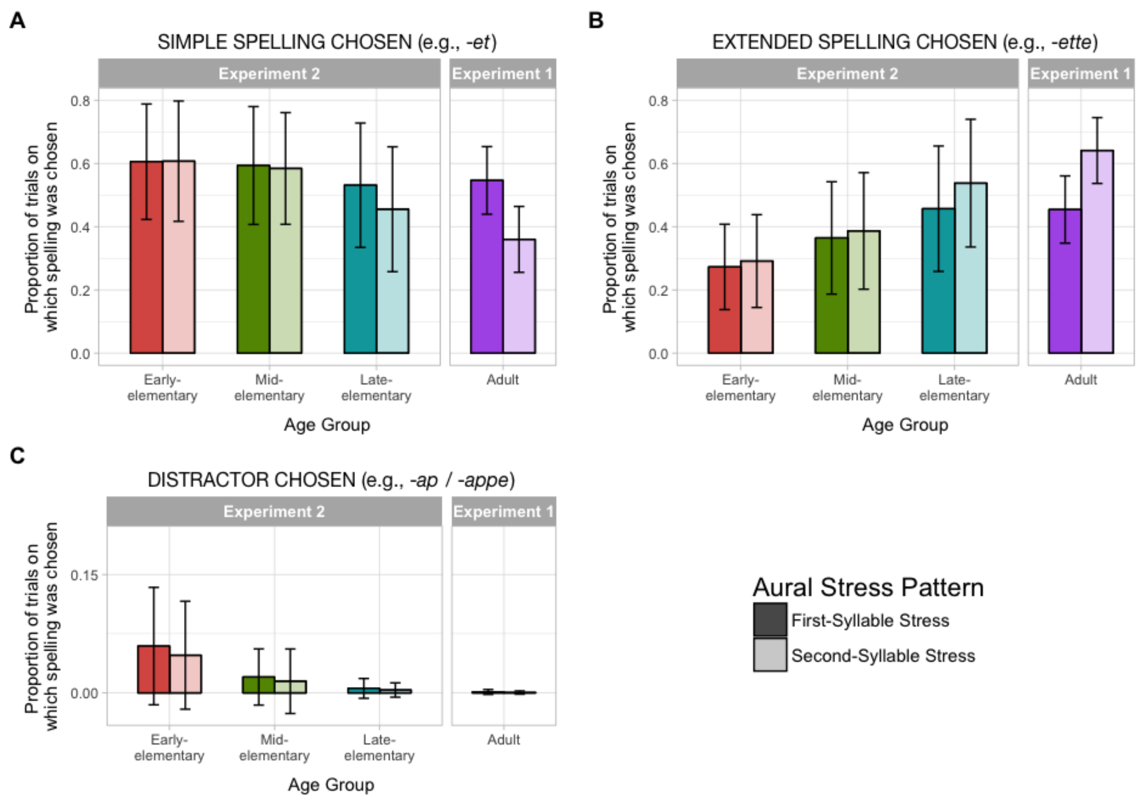


Figure 4.1. Mean rates at which participants chose the (A) simple spelling, (B) extended spelling, and (C) distractor spelling options as a function of age group and aural stress pattern. Error bars show standard deviations.

To determine whether adults' spelling choices are affected by lexical stress, we conducted two paired samples *t*-tests on the rates at which participants chose the simple and extended target spellings, respectively. Both analyses tested for whether these spelling rates differed as a function of an item's aurally presented stress pattern, and they are reported with Bonferroni-adjusted *p*-values that are corrected to reflect all tests in Experiment 1. The first analysis confirmed that participants were more likely to select the simple target spelling (e.g., *–et*) after hearing a pseudoword presented with first-syllable stress than one presented with second-syllable stress, $t(36) = 9.53$, $p < .001$, $d = 1.57$. The second showed that, conversely, participants were more likely to select the extended target spelling (e.g., *–ette*) after hearing a pseudoword presented with second-syllable than first-syllable stress, $t(36) = 9.50$, $p < .001$, $d = 1.56$. Figure 4.2 illustrates these effects, showing the paired differences in the rate at which participants selected each spelling after an item pronounced with first-syllable stress and the rate which they chose it after an item pronounced with second-syllable stress.

In both analyses, adults' spelling choices aligned with what we had expected based on the endings' associations with lexical stress in English. Our final analysis sought to determine whether participants chose the expected spellings more often than would be seen by chance.[12] To that end, we calculated the rate at which participants chose the expected spelling, averaging the selection of simple targets for items with first-syllable stress and the selection of extended targets for items with second-syllable stress ($M = 59.18\%$, $SD = 5.94\%$). A one-sample *t*-test confirmed that participants chose the expected target spelling at a significantly higher rate than would be expected by 50% chance, $t(36) = 9.43$, $p < .001$, $d = 1.55$.

---

[12] Because we found floor effects in distractor spelling selection, we used 50% chance, rather than 25%, as the standard against which to test participants' performance.

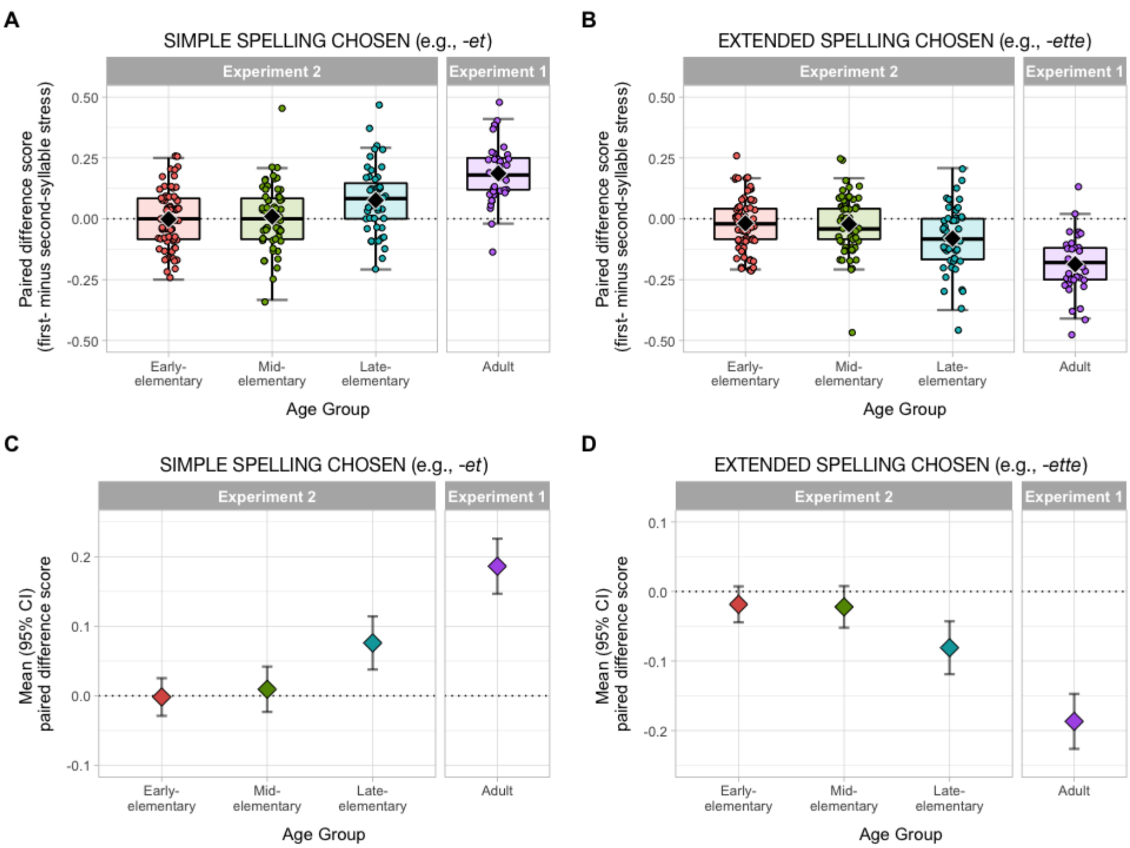


Figure 4.2. Paired difference scores in the rate at which participants chose simple and extended spellings after items pronounced with first- versus second-syllable stress. Panels A and B show scatterplots and boxplots of paired differences in the rates at which each participant chose the spelling options after items pronounced with first- and second-syllable stress. Panels C and D show the corresponding mean differences for each age group; error bars represent 95% confidence intervals.

These analyses converge on the conclusion that adult participants' spelling choices are influenced by the stress pattern with which pseudowords are pronounced. Participants preferentially chose the ending spellings that align most strongly with the presented stress pattern in real English words—simple spellings with first-syllable stress, and extended spellings with second-syllable stress—with large effect sizes in both cases. These results replicate Verrekia's (1996) finding that aural stress patterns affect adults' spelling choices. They are also consistent with related work showing that the spelling of word endings can influence adults' processing of stress in known English words (Kelly et al., 1998; Mundy & Carroll, 2013). Notably, our findings show that these prior results are unlikely to be accounted for by simple use of vowel quality, as this factor was controlled in our study. Thus, our findings suggest adults are indeed sensitive to the associations between word endings and lexical stress in English, and they use that sensitivity as a means of distinguishing between plausible spelling alternatives for a given sound. In Experiment 2, we consider whether—and when—the same might be true of young readers.

4.4 EXPERIMENT 2: DEVELOPING READERS

4.4.1 METHOD

Participants. We recruited 186 children across Grades 1 to 6 from elementary schools in an urban setting in eastern Canada. Data for 185 participants were included in analyses; one child did not follow task instructions and their data were therefore excluded. Participating children spanned the elementary school grades, with 25 in Grade 1 (mean age = 6.96 years, *SD* = 0.26), 47 in Grade 2 (mean age = 7.88 years, *SD* = 0.31), 37 in Grade 3 (mean age = 8.93 years, *SD* = 0.28), 25 in Grade 4 (mean age = 9.90 years, *SD* = 0.33), 30 in Grade 5 (mean age = 11.10 years, *SD* = 0.37), and 21 in Grade 6 (mean age = 11.96 years, *SD* = 0.35). Testing took place in the spring, toward the end of the school year. Children received a small token of appreciation (e.g., a pencil or eraser) as thanks for their participation.

Materials. The spelling choice task was similar to that used with adults in Experiment 1, modified slightly for use with young children. Participants listened to a set of 48 aurally-presented pseudowords, and for each, they were asked to select their preferred spelling from among four forced-choice options.

Pseudoword items. To reduce task length and ensure suitability for developing readers, we selected a subset of 6 endings from those used in Experiment 1: *–et*/*–ette*, *–um*/*–umb*, *–y*/*–ee*, *–o*/*–eau*, *–eak*/*–ique*, and *–us*/*–uss*. The chosen endings satisfied two criteria. First, we confirmed that our target endings' spellings appear in elementary texts, and that those spellings had our target pronunciation in most or all of words in which they appear (Zeno, Ivens, Millard, & Duvvuri, 1995). Secondly, we confirmed that the majority of disyllabic words with each target ending had our intended stress pattern (e.g., first-syllable stress for words ending with *–et*; second-syllable stress for words ending with *–ette*).

As in Experiment 1, we conducted acoustic analyses on the 48 pseudowords that were presented to children. Mean PVI scores for the items' pitch (Hz) and intensity (dB) are shown in Table 4.1. One-sample *t*-tests showed that items pronounced with first-syllable stress had PVI scores that were significantly greater than zero for both pitch, $t(47) = 15.67$, $p < .001$, $d = 2.26$, and intensity, $t(47) = 15.33$, $p < .001$, $d = 2.21$. Similarly, items pronounced with second-syllable stress had PVI scores significantly lower than zero for both pitch, $t(47) = -10.11$, $p < .001$, $d = -1.46$, and intensity, $t(47) = -9.46$, $p < .001$, $d = -1.36$. These analyses indicate that the pseudowords presented to children were produced with their intended stress patterns.

We also used the procedure described in Experiment 1 to confirm that vowel quality was sufficiently controlled across the pseudowords presented to children. Mean frequencies for F1, F2, and F3 are shown in Table 4.2 and were analyzed through a repeated-measures MANOVA with the formants as dependent variables (F1; F2; F3) and aural stress pattern (first-syllable stress; second-syllable stress) as a within-subjects factor. This analysis showed no effect of aural stress pattern on the combined vowel formants, $\Lambda = .98$, $F(3, 93) = 0.73$, $p = .54$, $\eta_p^2 = .02$. This helps to confirm that vowel quality was similar across our pseudoword stimuli.

Spelling choice task. We presented children with eight pseudowords for each of the six endings using the procedure described in Experiment 1. Participants therefore completed a total of 48 items, distributed in the printed booklets such that only one instance of each ending appeared per page.

Procedure. Participants were tested in groups, either in their classrooms or in a quiet room at their school. Throughout the session, participants were monitored by two experimenters to ensure that they were following instructions and could not see each other's responses. The aurally-presented items were arranged in a continuous playlist, as in Experiment 1, though the delay between items varied based on the pace at which children responded in each group. Participants were told that they would hear a series of made-up English words, and that for each, they should circle the spelling that best represented what they heard. Children were assured that the task had no right or wrong answers and were encouraged to simply choose the option that they saw as best. Two practice items were used to familiarize children with the task. The experimenter paused between items until all participants had made their selection and repeated the pseudowords as needed. In total, the testing session lasted approximately 20 minutes.

4.4.2 RESULTS AND DISCUSSION

To analyze children's spelling choices, we combined the participants into grade groupings that capture the early elementary (Grades 1 and 2; $n = 72$), mid elementary (Grades 3 and 4; $n = 62$), and late elementary school years (Grades 5 and 6; $n = 51$).

Doing so allows us to explore cross-sectional trends over the elementary school grades, and ensures sufficient power to detect our effects of interest.

The left-hand panels of Figure 4.1 show the mean rates at which participants chose the simple target, extended target, and distractor spellings, presented as a function of participant grade level and pseudoword aural stress pattern. When calculating spelling selection rates, we treated responses as missing if the participant either (a) did not select a spelling for a given item, or (b) selected more than one spelling for a given item.[13] Across the task, 0.2% of all data points were missing. And as Figure 4.1 shows, there were floor effects in children's selection of distractor spellings at all grade levels. Thus, even among children in the early elementary grades, spelling choices were largely restricted to the two phonemically plausible alternatives.

To determine whether children's spelling choices were affected by the stress pattern with which pseudowords were pronounced, we conducted separate 3 x 2 analyses of variance. Each included grade level (early-elementary, mid-elementary, late-elementary) as a between-subjects factor and aural stress pattern (first-syllable stress; second-syllable stress) as a within-subjects factor. The first analysis used the rate at which children chose the simple spelling option (e.g., *–et*) as its outcome variable, whereas the second used the rate at which children chose the extended spelling option (e.g., *–ette*). To avoid interpreting spurious effects, we report Bonferroni-adjusted *p*-values that correct for the total number of comparisons run across these two analyses.

Simple spelling choice. The first ANOVA, examining children's selection of simple target spellings, showed two main effects. The first was of grade level, $F(2, 182)$

[13] This step was not necessary in Experiment 1, as neither situation occurred in the adults' data.

= 6.66, $p = .002$, $\eta_p^2 = 0.068$. Children in the late-elementary grades chose simple target spellings less often ($M = 49.31\%$, $SD = 18.5$) than did children in the early-elementary ($M = 60.63\%$, $SD = 17.8$) or mid-elementary grades ($M = 58.88\%$, $SD = 17.0$; Bonferroni-adjusted *p*s < .03), while early- and mid-elementary students chose simple spellings equally often ($p = 1.00$). There was also a significant main effect of aural stress pattern, $F(1, 182) = 8.95$, $p = .003$, $\eta_p^2 = 0.047$, with slightly greater selection of simple spellings for pseudowords with first-syllable stress ($M = 58.08\%$, $SD = 19.0$) than second-syllable stress ($M = 55.75\%$, $SD = 19.8$).

These main effects were qualified by a significant interaction, $F(2, 182) = 6.31$, $p = .002$, $\eta_p^2 = 0.065$. To interpret the interaction, paired-samples *t*-tests (reported with Bonferroni-adjusted *p*-values[14]) separately compared the rate at which children in each grade level chose simple target spellings as a function of aural stress pattern. Participants in the late elementary grades chose simple spellings more often after hearing pseudowords with first-syllable stress than those with second-syllable stress, $t(50) = 4.00$, $p = .002$, $d = 0.561$. However, pseudowords' aural stress pattern had no impact on the rate at which children chose simple spellings in either the early elementary grades, $t(71) = -0.13$, $p = 1.00$, $d = -0.015$ or the mid elementary grades, $t(61) = 0.58$, $p = 1.00$, $d = 0.072$. Figures 4.2a and 4.2c illustrate these effects of aural stress pattern by grade group.

Extended spelling choice. The second ANOVA, examining children's selection of extended target spellings, showed a very similar pattern of results. The main effect of grade level was significant, $F(2, 182) = 26.46$, $p < .001$, $\eta_p^2 = 0.225$, with progressively

[14] Bonferroni corrections for these *t*-tests adjust for nine total comparisons: (a) simple spelling choice rates, (b) extended spelling choice rates, and (c) the expected target selection rate compared against 50% chance, each with comparisons at the early-, mid-, and late-elementary grade levels.

more selection of extended target spellings when moving from the early-elementary (M = 28.19%, SD = 13.0) to mid-elementary (M = 37.54%, SD = 17.1) to late-elementary grades (M = 49.71%, SD = 18.8), ps < .01. The main effect of aural stress pattern was also significant, $F(1, 182) = 20.471$, $p < .001$, $\eta_p^2 = 0.102$, with greater overall selection of extended spellings for pseudowords pronounced with second-syllable stress (M = 39.08%, SD = 20.1) than first-syllable stress (M = 35.42%, SD = 18.5).

Again, these main effects were qualified by a significant interaction, $F(2, 182) = 4.74$, $p = .01$, $\eta_p^2 = 0.05$. Paired-samples t-tests with Bonferroni-adjusted p-values separately compared the rate at which children in each grade level chose extended target spellings as a function of aural stress pattern. Participants in the late elementary grades chose extended target spellings more often after hearing pseudowords with second-syllable stress than those with first-syllable stress, $t(50) = 4.27$, $p < .001$, $d = 0.598$. However, pseudowords' aural stress pattern had no impact on the rate at which children chose extended spellings in either the early-elementary grades, $t(71) = 1.43$, $p = 1.00$, $d = 0.172$, or the mid-elementary grades, $t(61) = 1.48$, $p = 1.00$, $d = 0.188$.[15] Figures 4.2b and 4.2d illustrate these aural stress pattern by grade group effects.

The simple and extended spelling analyses both suggest that it is not until the late-elementary grades that children begin to use stress to guide their spelling choices. Indeed, the overall rate at which children in Grades 5 and 6 chose the predicted target spellings was significantly higher than 50% chance, albeit only modestly so (M = 53.43%, SD =

[15] Three of the six extended endings used (*–ette*, *–ique*, and *–eau*) are of French origin, which is notable given that our participants receive French instruction as part of their core curriculum starting in Grade 4. To address the possibility that our effects were driven by the three French endings, we reanalyzed children's extended spelling choices after excluding the items that contained those endings. Doing so did not change the pattern or magnitude of results.

6.99), $t(50) = 3.50$, $p = .009$, $d = 0.49$. Although the late-elementary students' pattern of results was similar to those of our adult participants, the effects seen in adults were quite a bit stronger.[16] This suggests that while we may have identified a point at which children first begin to use stress to guide their spelling choices, we have not captured the point at which they do so to the same extent as experienced adult readers. It may be the case that children's sensitivity to the links between word endings and lexical stress continues to develop through adolescence.

On the other hand, children in the early- and mid-elementary school grades showed no signs of using stress to inform their spelling choices. In fact, children in Grades 1 and 2 chose the expected target spellings at rates significantly *lower* than 50% chance level ($M = 44.80\%$, $SD = 9.04$), $t(71) = -4.86$, $p < .001$, $d = -0.57$, reflecting their overall preference for more-frequent simple target spellings regardless of the aural stress manipulation. The rate at which children in Grades 3 and 4 chose expected target spellings was at chance level ($M = 48.98\%$, $SD = 7.49$), $t(61) = -1.06$, $p = 1.00$, $d = -0.13$. In other words, although children in the mid-elementary grades were more inclined to choose extended spellings than their younger peers, they did not yet use pseudoword stress patterns to systematically guide their spelling choices.

4.5 GENERAL DISCUSSION

When spelling in English, it is often the case that a given sound sequence can be represented in multiple ways (Kessler & Treiman, 2001; Ziegler et al., 1997). Fortunately for spellers of English, there are probabilistic regularities in the writing system that can

[16] To make the effects more directly comparable, we reanalyzed adult participants' data using only the subset of endings from Experiment 2. Doing so did not change the pattern or magnitude of effects reported in Experiment 1.

help them to decide between the alternatives, and an individual's sensitivity to those regularities may aid in making spelling decisions (e.g., Hayes et al., 2006; Treiman & Kessler, 2006). The current study examined the effects of lexical stress as one such regularity, extending our understanding of spelling decisions to capture a key feature of multisyllabic English words. In a forced-choice spelling task that disentangled stress and vowel quality, we showed that English-speaking adults used aurally-presented stress patterns to guide their preferred spelling of disyllabic pseudowords. Similarly, by Grades 5 and 6, children used stress to inform their spelling decisions. On the other hand, younger children did not do so, tending to prefer simple spellings (e.g., *–et*) over extended spellings (e.g., *–ette*) of pseudowords, regardless of the presented stress pattern.

4.5.1 READERS' USE OF THE ASSOCIATIONS BETWEEN ENGLISH STRESS AND SPELLING

Our findings in adults add important corroborating evidence to the existing research on readers' use of stress when processing written English words. Several previous studies have implicated stress in visual word recognition and naming of known words (Kelly et al., 1998; Mundy & Carroll, 2013), and have suggested that readers use word endings when assigning stress to unfamiliar words (Arciuli & Cupples, 2006; Arciuli et al., 2010). Yet, as noted earlier, these studies were not able to fully separate the effect of lexical stress from the effect of vowel quality. The spelling choice paradigm used in our study addressed this confound by controlling vowel quality across pseudowords with different stress patterns. And encouragingly, our results converge with those of prior work, which helps to increase our confidence in the interpretations that have been given in the literature to date. That is, adult readers do indeed seem to be sensitive to the associations between ending spellings and lexical stress in English, and that sensitivity affects their processing of written words.

On the other hand, our findings in children suggest that it takes a fair bit of time for readers to learn about and fully use stress regularities to guide their spelling selections. The most straightforward support for this claim comes from the fact that children in the early- and mid-elementary grades did not use stress to choose among spelling alternatives. Instead, regardless of a word's aural stress pattern, our younger participants preferentially chose the simple spelling option (e.g., *–et*). In contrast, we saw that children in the late-elementary grades had begun to use stress to inform their spelling choices, preferentially choosing simple spellings for items presented with first-syllable stress and extended spellings for those presented with second-syllable stress. This points to an emerging sensitivity to the links between stress and spelling.

Yet even then, children in the late-elementary grades may not have been taking full advantage of lexical stress as a cue to spelling—a possibility that gets support from a comparison of our findings across experiments. Specifically, the effect of stress on spelling in older children was fairly modest when compared to the larger effects seen in adults (recall that the approximate *d*s were 0.6 and 1.6, respectively). In other words, although we may have identified the approximate point at which children first use stress to guide spelling choices, it seems that we have not identified the point at which they do so in an adult-like way. Bearing in mind that our conclusions are drawn from a cross-sectional study, this suggests that children's sensitivity to the associations between ending spellings and stress may continue to develop beyond the elementary grades. Very little work to date has explored this sensitivity in older children. In one existing study, when adolescents (ages 13 to 17) were asked to place stress in written pseudowords, they assigned first-syllable stress more often to items whose endings cued first-syllable stress than those that cued second-syllable stress (Arciuli & Paul, 2012). This effect was large and was comparable to that seen in adults completing the same task (Arciuli & Cupples, 2006), though the data did not allow for a precise examination of differences across adolescence. By extending our spelling choice paradigm into older groups of children, we may get a more complete picture of spellers' emerging use of stress to distinguish between plausible alternatives.

The developmental trend that we saw in our study should be interpreted with reference to the small body of existing research on children's sensitivity to ending spellings as stress cues. In the current study, the first signs of stress being used to guide spelling choices emerged in our oldest group of children (ages 10 and 11 years). This is slightly later than the point at which similar effects have been seen in prior work: when pronouncing written pseudowords, children have shown clear sensitivity to the links between spelling and stress by age 9 (Arciuli et al., 2010; Colombo et al., 2014; Sulpizio & Colombo, 2013). In contrast, children of the same age in our study showed no such sensitivity in their spelling decisions (recall that participants in the mid-elementary group were approximately 9 years old). It is possible that the mid-elementary aged children in Arciuli and colleagues' (2010) study used vowel quality to drive their pronunciations, resulting in a discrepancy with the findings of our study when vowel quality was removed as a potential cue. With that said, we have no strong reason to expect a shift from the use of vowel quality in the mid-elementary grades to stress in the late-elementary years. Instead, a likely explanation for the discrepancy lies in the word endings used in the two studies. In particular, the endings used by Arciuli and colleagues

(2010) are more prevalent in children's earliest texts (e.g., *–et*, *–oon*, *–act*) than our own extended spellings (e.g., *–ette* and *–umb*), according to a corpus of children's texts (Zeno et al., 1995). As such, children may have had fewer learning opportunities for our endings, leading, in turn, to a slight delay in the use of stress to guide spelling decisions. Of course, this more limited exposure to extended spellings also helps to explain the younger children's general preference for simple spellings.

4.5.2 LEXICAL STRESS AND MODELS OF SPELLING CHOICE

Our findings speak against views of the spelling process which claim that experienced spellers choose among spelling alternatives solely in line with how frequently the different spellings appear in the language (e.g., Barry & Seymour, 1988). This view may characterize how children in our youngest age groups approached the pseudoword spellings, as the simple spellings that they preferentially chose are indeed more common than their extended spelling alternatives. Yet by the late elementary grades and in adulthood, we saw no overall preference for simple spellings. Instead, after hearing a pseudoword pronounced with second-syllable stress, older children and adults preferred the less frequent extended spellings over the more frequent simple ones. This preference follows the statistical associations that exist between the extended spellings and second-syllable stress in English, reflecting our older participants' sensitivity to word endings as lexical stress cues. These more experienced spellers' performance aligns with the view that, when making spelling choices, spellers draw on factors that constrain the likelihood of a sound-to-spelling correspondence being used (e.g., Hayes et al., 2006; Treiman & Kessler, 2006).

Although the processes involved in spelling have not been modelled as extensively as those of reading, Houghton and Zorzi (2003) provide a well-developed model for the spelling of monosyllabic words. Their model takes seriously the notion that spelling output can be influenced by phonological features of the word in which a sound appears. It does so using a dual-route connectionist structure, with its phonological route consisting of a two-layer associative network capable of learning sublexical sound-to-spelling regularities through corpus training. Critically, all input (phoneme) units in the network are connected to all output (grapheme) units, which allows the model to simulate the effects of surrounding phonological context on spelling (e.g., Hayes et al., 2006; Treiman, Kessler, & Bick, 2002). This structure represents an important step in modelling how spellers choose between multiple plausible alternatives.

With that being said, Houghton and Zorzi's (2003) model is limited to monosyllabic words, and thus it has no mechanism by which to incorporate information about lexical stress into spelling output. Our results would suggest that an effective model of spelling should be able to do so. Recently, structurally-similar models of reading aloud have had success in modelling various effects of lexical stress in English (Perry, Ziegler, & Zorzi, 2010), which might indicate that there is scope for modifying Houghton and Zorzi's (2003) framework to account for our findings. Regardless, further work is needed to ensure that models of the spelling process incorporate the features of multisyllabic words that affect spelling decisions.

4.5.3 PRACTICAL IMPLICATIONS

From a practical standpoint, the fact that stress can be used to distinguish between spelling alternatives is encouraging: stress adds a degree of regularity to the English writing system, which is often regarded as inconsistent. With that said, our results show that children's use of stress to guide spelling is fairly late to emerge and is initially modest. In other words, although stress can be used to benefit spelling when faced with more than one spelling alternative, young children may not take full advantage of that benefit. Our findings in adults suggest that spellers can eventually do so, if given enough time and experience with the regularities in English text. Yet there may be instructional opportunities to alert younger students to the ways in which spelling relates to stress. To date, one intervention study has shown that typically-developing children can be explicitly taught the associations between spelling and stress in English (van Rees, Ballard, McCabe, Macdonald-D'Silva, & Arciuli, 2012). While promising, that study involved multiple hours of individual training sessions—an instructional time demand that likely outweighs any benefit to children's spelling.

As an alternative form of instruction, others have advocated for a less direct approach to teaching spelling regularities. For example, Treiman and Kessler (2006) suggest that teachers draw children's attention to regularities that can be helpful in selecting between different spellings of the same sound, without explicitly teaching each instance of those regularities. Doing so, and then creating opportunities for students to encounter those regularities in text (for example, through exposure to a variety of words that end in *–et* and *–ette*) may help students to more readily use stress to guide spelling. Based on our results, this type of instruction might be particularly helpful in the mid-elementary grades, when students increasingly recognized that extended spellings were plausible alternatives to simple spellings, but did not yet use stress to differentiate between spelling options.

4.5.4 LIMITATIONS AND FUTURE DIRECTIONS

When designing this study, we focused on the statistical associations that exist between our ending spellings and lexical stress patterns, arguing that participants draw on these associations when using stress to inform their spelling choices. However, the endings associated with second-syllable stress were largely extended spellings that contained silent letters, whereas endings associated with first-syllable stress were simple spellings that did not (per Kelly et al., 1998). As such, our findings cannot distinguish between these two characteristics of our ending pairs. With that caveat in mind, we argue that participants' use of stress to guide their spelling choices is more likely to draw on the endings' associations with stress than on the presence or absence of an extended spelling. For one thing, corpus analyses suggest that linking stress placement with extended spellings would be ineffective as a broad strategy. As an example, endings with doubled consonants are not systematically associated with second-syllable stress (unless they also end with a silent *e*; Arciuli & Cupples, 2006), making the presence of additional letters themselves an unreliable cue. Furthermore, our findings generally converge with those of studies in which the two ending features were not conflated (e.g., Arciuli et al., 2010). These points give us some confidence in our interpretation, though future work that includes a wider variety of phonemically-similar ending pairs will be needed to fully disentangle the two possibilities.

Another consideration for future work relates to statistical analysis. In the current study, we used analysis of variance (ANOVA) and paired-sample *t*-tests to identify differences in the rate at which participants chose a given spelling after hearing pseudowords pronounced with first- versus second-syllable stress. Doing so yielded a promising pattern of results and allowed us to report widely-known measures of effect size (partial eta squared and Cohen's *d*), both of which speak to cross-sectional trends in readers' use of stress to guide spelling decisions. However, there are issues with the use of ANOVA when an outcome variable is inherently categorical. For example, the confidence intervals derived from these ANOVAs can yield implausible values (i.e., proportion rates greater than 1 or less than 0), and binomial data is prone to heterogeneity of variance, violating a key assumption of ANOVA (e.g., Jaeger, 2008). Given that our data resulted from a forced-choice task in which participants' responses were largely restricted to two plausible spelling options (e.g., *–et* or *–ette*), mixed-effects logistic regression may represent a more appropriate form of analysis in future research.

Our use of a forced-choice spelling task was critical to addressing our research question, in that it gave us a high degree of control over the ending spellings that were available to participants. It allowed us to focus on spellings with confirmed associations with lexical stress in English. A spelling production task would have lacked this control, as participant-generated spellings might include endings that are not used in real English words or that are unrelated to lexical stress. This has been the case in prior generative spelling research (Verrekia, 1996). Despite this strength, the forced-choice format limits the conclusions that we can draw about the extent to which people use stress to generate their own spellings. To complement the current study, it will be important for future research to determine whether our findings extend to a more naturalistic spelling context, particularly in children.

In sum, the current study makes both empirical and theoretical contributions. It confirms prior evidence that adult readers are sensitive to the links between spelling and stress in English (e.g., Kelly et al., 1998; Verrekia, 1996), and it extends this line of research with children (e.g., Arciuli et al., 2010). Importantly, it does so with a method that eliminates vowel quality as a potential confound. This implicates lexical stress—and not its main phonemic correlate—as the source of our findings. The findings have implications for theories of spelling (Barry & Seymour, 1988; Houghton & Zorzi, 2003), which to date have focused on spelling regularities at the phoneme-to-grapheme level. Specifically, we show that young children tend to preferentially choose the more frequent spelling alternatives, regardless of a word's aural stress pattern. In contrast, adults and children in the late-elementary grades use lexical stress to guide their spelling choices. Experienced spellers thus appear to be sensitive to fine-grained spelling regularities and show no broad preference for the more frequent spelling alternatives. Taken together, our findings reinforce the claim that lexical stress is an important component of processing multisyllabic words, which can be applied to aid the often challenging task of choosing among plausible spellings in English.

4.6 APPENDIX F: PSEUDOWORD ENDING SPELLINGS

Characteristics of the 14 endings used in the Spelling Choice task completed by adults. Endings marked with asterisks were also used in the task completed by children.

		Simple spelling			Extended spelling		
Ending sound		Ending spelling	Proportion of words with 1st-syllable stress	Type frequency of spelling [a]	Ending spelling	Proportion of words with 2nd-syllable stress	Type frequency of spelling [a]
/ɑr/		*–ar*	0.88	89	*–arre*	1.00	1
/ik/	*	*–eak*	0.78	24	*–ique*	1.00	12
/in/		*–ene*	0.70	12	*–een*	0.79	32
/ɛss/		*–es*	0.98	42	*–esse*	1.00	3
/ɛsk/		*–esk*	1.00	1	*–esque*	1.00	3
/ɛt/	*	*–et*	0.89	216	*–ette*	0.85	20
/oʊ/	*	*–o*	0.90	181	*–eau*	0.56	10
/oʊr/		*–or*	0.94	70	*–ore*	0.56	45
/ɒs/		*–os*	1.00	13	*–oss*	0.75	15
/ɒt/		*–ot*	0.95	84	*–otte*	0.80	5
/u/		*–u*	0.72	23	*–oo*	0.71	23
/ʌm/	*	*–um*	1.00	58	*–umb*	0.75	9
/ʌs/	*	*–us*	0.97	73	*–uss*	0.67	8
/i/	*	*–y*	0.98	1147	*–ee*	0.54	77

[a] Type frequencies were calculated across mono- and disyllabic English words, per Barry and Seymour (1988)

CHAPTER 5
DO ORTHOGRAPHIC CUES TO STRESS AFFECT READING IN CONNECTED TEXT?

5.1 ABSTRACT

In English, written word endings act as probabilistic cues to a word's lexical stress pattern. Readers are sensitive to these statistical associations between stress and spelling, using endings to guide stress placement when reading isolated words. However, we do not yet know if readers' use of endings as stress cues extends to the processing of words in connected text. In the current study, we explored this question in adult readers ($N = 53$) and children in Grades 3 to 6 ($N = 101$). Participants read texts for comprehension while circling a specified target letter. This letter-detection task can speak to the processes involved in reading words, as readers detect fewer letters when lexical access is relatively quick. Interestingly, adults detected fewer target letters in words whose endings accurately cued their stress patterns than in words whose endings gave misleading cues to stress ($d = 0.41$). This suggests that adult readers draw on written word endings as cues to stress during the everyday task of reading for comprehension. However, word endings had no effect on children's letter detection rates. Together, these findings clarify the scope of word endings' role as cues to stress in English.

5.2 INTRODUCTION

Traditionally, investigations of reading and visual word recognition have focused on characteristics of monosyllabic words (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). While important, this work leaves a gap in our understanding of the multisyllabic words that make up over 90% of the English lexicon (Baayen, Piepenbrock,

& van Rijn, 1995). Lexical stress—which refers to the pattern of relative emphasis across syllables in a word—is a phonological characteristic that does not apply to monosyllables, yet is a key feature of multisyllabic words. For example, correct stress placement helps to distinguish between word meanings (e.g., the noun *REcord* and the verb *reCORD*) and separates real English words from pronounceable nonwords (e.g., **soLItary*). And although it is integral to a word's phonology, lexical stress is not marked directly in written English. In recent years, substantial gains have been made in our understanding of lexical stress—both the role it plays in visual word recognition (e.g., Kelly, Morris, & Verrekia, 1998), and its activation when reading for comprehension (e.g., Ashby & Clifton, 2005; Breen & Clifton, 2011). A fruitful area of this research has focused on the role that written word endings play as cues to stress placement (e.g., the ending *–et* is associated with first-syllable stress, as in the word *ASSet*; Arciuli & Cupples, 2006). By mid-childhood, and into adulthood, readers recognize and use statistical links between the spelling of word endings and lexical stress patterns (e.g., Arciuli, Monaghan, & Ševa, 2010; Kelly et al., 1998). To date, these findings have emerged in tasks where participants read isolated words and pseudowords in controlled experimental tasks. As a result, the extent to which readers use word endings as stress cues in more naturalistic reading situations is not yet clear. In the current study, we explore whether written word endings act as cues stress when reading connected text for comprehension.

5.2.1 ENDING SPELLINGS AS LEXICAL STRESS CUES

When readers encounter multisyllabic written words in English, one piece of information they can use to establish stress placement is the spelling of the word's ending. Corpus analyses have shown that many ending spellings are statistically associated with specific stress patterns (Arciuli & Cupples, 2006; Arciuli et al., 2010; Kelly et al., 1998; Monaghan, Arciuli, & Ševa, 2016). For example, approximately 90% of English words ending in *–et* have first-syllable stress (e.g., *COMet*; *ASSet*). As a result, the spelling *–et* acts as a reasonably strong, though probabilistic, cue to the word's stress pattern. Other endings have somewhat weaker links with stress, yet are still more associated with one stress pattern than another. The ending *–ice*, for instance, is 78% associated with first-syllable stress (*SERvice*; *CREvice*)—a word with that ending is more likely than not to have first-syllable stress. Notably, these associations exist in English texts read by adults (e.g., Arciuli & Cupples, 2006), as well as those read by young children (Arciuli et al., 2010). As such, word endings are available as potential stress cues from early in children's reading development.

The behavioural research to date suggests that readers are sensitive to the associations between ending spellings and lexical stress. Much of this work has focused on the stress patterns that readers apply to unfamiliar pseudowords. For example, when presented with disyllabic written pseudowords and asked which syllable should receive emphasis, English-speaking adults preferentially choose the stress pattern indicated by the word's ending (Arciuli & Cupples, 2006; see e.g., Colombo, Deguchi, & Boureux, 2014 for comparable findings in Italian-speaking adults). Similar results have been reliably found in children by around 9 years of age. In one study, for instance, 9- to 12-year-old children tended to pronounce written pseudowords with stress patterns that aligned with the patterns predicted by their endings. The study's 5- to 8-year-old participants made some use of endings as cues to stress, producing first-syllable stress slightly more often when the pseudoword's ending suggested first-syllable stress than second-syllable stress. However, before age 9 these children made only limited use of endings, as they assigned first-syllable stress to a majority of words, including those whose endings suggested second-syllable stress. In other words, the tendency to use endings to guide stress increased with age; endings had a stronger effect on stress placement in older children than their younger peers (Arciuli et al., 2010; see also Sulpizio & Colombo, 2013 for similar findings in Italian). This work has helped to establish that the statistical links between ending spellings and lexical stress in English have behavioural impacts on both skilled and developing readers.

Interestingly, endings also seem to cue stress placement in familiar written words. A study by Kelly and colleagues (1998) provides a seminal illustration of this point. In that study, the researchers identified a set of word endings that cued either first- or second-syllable stress, and for each, selected words whose stress pattern was either consistent or inconsistent with that ending. Consistent items had ending spellings that were associated with the correct stress pattern for the word in question. For example, the endings of *COMet* and *rouLETTE* are associated with first- and second-syllable stress, respectively, making their ending spellings consistent with their stress patterns. On the other hand, inconsistent items had different stress patterns than would be predicted by their endings (e.g., *caDET* and *PALette*). English-speaking adults completed either a naming or visual lexical decision task with these items, and in both cases responded more quickly and accurately to words whose endings were consistent with lexical stress than those whose endings were inconsistent. This finding has been replicated in English-speaking adults (Mundy & Carroll, 2013), and aligns with research done in other languages (e.g., Burani & Arduino, 2004; Burani, Paizi, & Sulpizio, 2014).

Kelly and colleagues (1998) interpreted their findings by proposing that inconsistent, misleading cues to stress impede lexical access of written words. Specifically, their interpretation suggests that inconsistent ending cues lead to an increase in stress placement errors, at least temporarily. For example, a reader presented with the word *cadet* might erroneously read it as the nonword /'KÆD ɪt/. The time it takes to correct this error results in slower lexical access, while a failure to correct it yields an inaccurate response. In contrast, word endings that are consistent with their word's stress pattern avoid this problem—these ending accurately cue stress, reducing the likelihood of a stress placement error (e.g., /kə 'MƐT/ for *comet*) and allowing for quicker, more accurate lexical access.

Importantly, though, research to date on the behavioural effects of word endings as stress cues has focused on words that are read in isolation. This is true of all empirical studies described above (Arciuli & Cupples, 2006; Arciuli et al., 2010; Colombo et al., 2014; Kelly et al., 1998; Mundy & Carroll, 2013). It is also the case for efforts to incorporate ending spellings into computational models of multisyllabic word reading (Perry, Ziegler, & Zorzi, 2010; Rastle & Coltheart, 2000; Ševa, Monaghan, & Arciuli, 2009). These highly controlled isolated word reading tasks have been valuable in establishing that ending spellings act as cues to lexical stress in written English words. However, isolated word reading is quite removed from the everyday task of reading texts for comprehension, which is how readers typically encounter written words. This disconnect leaves the scope of word endings' effects somewhat unclear. Specifically, we do not yet know whether—or to what extent—ending cues affect stress placement and lexical access when reading words in connected text. In the current study, we aim to address this gap.

The need to test whether word endings serve as stress cues in connected texts becomes particularly clear when we consider the unique demands of tasks used to explore isolated word reading and recognition. Lexical decision tasks serve as a good example. Unlike when reading for comprehension, the goal when making lexical decisions is to differentiate real words from nonwords—a process that is measurably affected by the presence and characteristics of those nonwords (e.g., Stone & Van Orden, 1993; Yang & Zevin, 2014; Yap, Balota, Cortese, & Watson, 2006). There is even evidence to suggest that nonwords' characteristics specifically affect the extent to which readers use word endings as cues to stress. A recent study of Italian-speaking adults (Colombo & Sulpizio, 2015) examined whether word endings' consistency with stress patterns would affect lexical decision responses. Importantly, the only sign of an ending consistency effect came from a version of the task in which nonword spellings overlapped with the spellings of the real-word items, requiring in-depth orthographic processing in order to successfully respond. In a lower-demand version of the task where words and nonwords were more distinct, word endings had no effect on lexical decision speed or accuracy (Colombo & Sulpizio, 2015).

Given that these differing trends occurred within slight variations of the same task, it would be premature to conclude that the findings from work on isolated word reading can be applied to the task of reading connected text. As such, research that uses more ecologically valid reading tasks is needed to address the possibility that written word endings act as cues stress when reading for comprehension. In the current study, we explored this possibility by presenting participants with connected texts, which we asked them to read for comprehension. Within the texts were two sets of words with endings that were either consistent or inconsistent with their stress patterns. By taking this approach, we were able to maintain the controlled item comparisons typical of work on isolated word recognition (Kelly et al., 1998; Mundy & Carroll, 2013) while applying them in a task that more closely resembles everyday reading.

5.2.2 LEXICAL STRESS AND SILENT READING

Our study investigates whether readers use ending spellings as cues to stress while reading texts for comprehension. We explore the possibility that target words embedded in those texts will be processed more efficiently when their stress patterns are consistent with their endings than when their stress patterns are inconsistent with their endings. To our knowledge, this question has never been examined in tasks that involve reading connected texts. However, there is research to support our use of connected texts when looking at factors related to lexical stress. Evidence suggests that readers activate prosodic information—including representations of lexical stress—during silent reading for comprehension (e.g., Ashby & Clifton, 2005).

An early demonstration that stress is activated during silent reading comes from Ashby and Clifton (2005), who tracked participants' eye movements while they read sentences containing target words with differing numbers of stressed syllables. They found longer gaze durations for words with two stressed syllables (e.g., *FUNdaMENtal*) than words with only one (e.g., *sigNIFicant*), which they interpreted as suggesting that readers activate stress patterns in the late stages of lexical access. Building on this, work by Breen and Clifton (2011) has shown that silent reading is disrupted when a word's stress pattern violates the pattern readers have been cued to expect based on metrical structure or grammatical category. In their experiment demonstrating the latter, Breen and Clifton (2011) constructed sentences with noun–verb pairs, including some that involved stress changes (e.g., *ABstract* vs. *abSTRACT*) and some that did not (e.g., *rePORT*). Eye movements were disrupted when the sentence context forced readers to revise their expectations about a word's grammatical category. And notably, that disruption was strongest when the critical word's stress pattern had to be changed. Their findings support the claim that stress patterns are included in readers' representations of words in text, and that correcting misplaced stress is a measurable part of the reading process.

A parallel body of research suggests that readers process the letters in stressed syllables differently than unstressed ones while reading texts. For instance, in a series of proofreading experiments, Harris and Perfetti (2016) showed that adults detect spelling errors more reliably if those errors appear in stressed syllables than unstressed ones, at least when text comprehension was among participants' goals. That is, when adults were instructed to both detect errors and read for comprehension, and when they were instructed to read for comprehension alone (with an incidental request to mark any errors that they happened to notice), they detected more errors in stressed than unstressed syllables. These findings align with early research into the effects of stress on letter detection (Drewnowski & Healy, 1982; Goldman & Healy, 1985). When reading passages for comprehension and simultaneously searching for a specified target letter, adults were better able to detect letters that occurred in the stressed syllables of high frequency words than letters that appeared in unstressed syllables (though not when targets appeared in a word's first syllable; Drewnowski & Healy, 1982). These findings indicate that syllable stress affects letter detection during reading for comprehension. However, they do not address whether word endings act as cues to lexical stress when reading for comprehension. This is the question that we explored in the current study.

5.2.3 THE CURRENT STUDY

To explore whether readers use endings as stress cues while reading connected text, our study adopts the letter-detection paradigm used in the research described above (Drewnowski & Healy, 1982; Goldman & Healy, 1985). In this task, participants are asked to read a text for comprehension while also circling a specified target letter each time they notice it. This simple paper-and-pencil task has traditionally been used to detect what is known as the missing letter effect, wherein participants tend to systematically miss letters that occur in function words (vs. content words) and in frequent words (vs. infrequent words; e.g., Roy-Charland & Saint-Aubin, 2006; Saint-Aubin & Poirier, 1997). By examining this type of systematic letter omission pattern, researchers can use the missing letter effect task as a simple and flexible way to examine the psycholinguistic processes involved in reading words in connected text (Klein & Saint-Aubin, 2016).

Of interest for our purposes, the missing letter effect has been linked to the process of lexical access. Several theoretical accounts have been put forward to explain the effect (and they do so with varying degrees of success; see e.g., Klein & Saint-Aubin, 2016 for a review). Notably, a common thread among many of those accounts is that letter detection rates depend—at least in part—on the time spent processing a word (Greenberg, Healy, Koriat, & Kreiner, 2004; Moravcsik & Healy, 1995; Roy-Charland, Saint-Aubin, Klein, & Lawrence, 2007; cf. Koriat & Greenberg, 1991 for an exception). For instance, the attentional-disengagement model offers an integrative and empirically well-supported account for the patterns that have been observed in readers' letter detection rates (Roy-Charland et al., 2007). This model argues that information about the presence of target letters builds passively while a reader's attention is held on a given word, but decays gradually once that attention has been removed. As a result, readers are less likely to detect target letters in words from which attention is disengaged quickly. Importantly, then, characteristics of a word that influence the speed of lexical access should affect the rate at which readers detect the letters within words (Klein & Saint-Aubin, 2016).

Drawing on this foundation, the current study used a letter detection task to explore the use of word endings as cues to stress when reading connected text. We based our study's design and predictions on the claim that, in isolated words, inconsistent (and therefore misleading) ending cues impede readers' lexical access whereas consistent cues facilitate it (Kelly et al., 1998). If this is true while reading words in connected text, we would expect to see a higher target letter omission rate in our consistent ending condition (i.e., words whose endings are consistent with their stress patterns) than our inconsistent ending condition (i.e., words whose endings are inconsistent with their stress patterns). Our study tests this hypothesis in two experiments: one that focuses on adult readers, and one that focuses on elementary school-aged children.

In Experiment 1, our participants were skilled adult readers. Evidence to date on isolated word naming and recognition suggests that adults are sensitive to the links between word endings and lexical stress, and that they use that sensitivity when processing written words in English (Arciuli & Cupples, 2006; Kelly et al., 1998; Mundy & Carroll, 2013). Based on that, we expect our ending consistency manipulation to affect adults' letter detection rates, with more omissions in words whose endings are consistent with their stress pattern than in words whose endings are inconsistent. However, it is also possible that presenting words in connected text might diminish—or altogether negate—the ending consistency effects seen in prior research. If this is the case, we might see no difference in adults' letter omission rates across conditions.

Experiment 2 uses the same missing letter effect paradigm with children in the mid- to late elementary grades. In doing so, we build on prior empirical work showing that children use ending spellings to guide stress placement when reading unfamiliar pseudowords (Arciuli et al., 2010). As noted earlier, this prior work demonstrates that children can make some use of endings to guide pseudoword stress placement from as early as 5 years of age. Notably, though, it also shows that the effect of endings is stronger among 9- to 12-year-olds than among younger children. Our study therefore focuses on children in Grades 3 to 6. This age range is useful to our research goals for other reasons, as well. First is the fact that children encounter a rapidly increasing number of multisyllabic words during this period of their schooling (the number of encounters grows by 19,000 each year through the fifth grade, according to some estimates; Kearns et al., 2016). As a result, multisyllabic words—which always carry lexical stress—play a growing role in children's reading experience across the elementary years. At the same time, children's reading instruction and activities focus heavily on reading for meaning during the mid- and late elementary grades (Chall, 1983). This makes Grades 3 to 6 a useful range to focus on in a task that centrally involves reading for comprehension.

We do not expect that our youngest (Grade 3) participants will use endings as stress cues while reading connected text. In the prior research on pseudoword stress placement, children did not reliably pronounce words in accordance with their endings until around age 9 (Arciuli et al., 2010). Coincidentally, 9 years is the same approximate age at which children first show a missing letter effect based on word function (for content and function words of similar frequency; Saint-Aubin, Klein, & Landry, 2005). All told, this evidence suggests that Grade 4 is the earliest point at which we are likely to see an effect of ending cue consistency on readers' letter detection. Thus, while we are unlikely to see an effect of ending consistency in our Grade 3 participants, including these young readers increases the chance that we will be able to identify the effect's earliest emergence.

Extending Experiment 2's age range up to Grade 6 lets us explore whether the use of endings to cue stress changes across the upper elementary grades. In Arciuli and colleagues' (2010) study of pseudoword naming, endings' effect on stress placement increased slightly with age. As such, we might expect any effect of ending cues on letter detection to become larger with grade level. For instance, we might see a small effect in Grade 4 students, a moderate one in Grade 5 students, and a relatively larger effect in Grade 6 students. At the same time, it is certainly possible that reading words in connected text will reduce the importance of ending spellings for children, such that we do not see an effect of ending cue consistency until later in childhood, if at all.

5.3 EXPERIMENT 1: ADULT READERS

5.3.1 METHOD

Participants. Participants were 53 undergraduate university students recruited in eastern Canada (mean age = 21.15 years, *SD* = 2.65; 37 women). As compensation, participants earned either $5 or 0.5 percentage points toward a university course.

Materials. Three texts were constructed for this experiment, with an average passage length of 664 words. Passages were assigned a target letter, indicated to participants at the top of every page. The target letters were embedded in a set of critical words that we used to test for a missing letter effect. Table 5.1 summarizes the characteristics of these texts: their length, text complexity, designated target letter, and the number of target letters that appeared in critical and non-critical words.

Item selection. All of our critical words were disyllabic nouns pronounced with first-syllable stress. Words that can be pronounced with either first- or second-syllable stress (e.g., *suspect*) were excluded from consideration. With stress held constant, our manipulation rested on differences in the word endings' associations with stress. Thus, half of our critical words had endings that were primarily associated with first-syllable stress (consistent condition), whereas half had endings that were primarily associated with second-syllable stress (inconsistent condition).

To create our two conditions, we selected endings in pairs that had similarly strong associations with first-syllable stress (consistent condition) and second-syllable stress (inconsistent condition). For example, 78% of words ending with *–ice* have first-syllable stress, whereas its paired ending *–uce* appears in words with second-syllable stress 72% of the time. All associations between endings and stress were calculated from among the disyllabic items in the CELEX English lexical database (Baayen et al., 1995). Confirming our manipulation, associations with first-syllable stress were significantly stronger for endings in the consistent condition (*M* = .77, *SD* = .07) than for endings in the inconsistent condition (*M* = .30, *SD* = .09), $t(8) = 9.20$, $p < .001$.

The target letters used to test for the missing letter effect appeared in these word endings. As such, the endings in a pair shared a letter in the same position (e.g., *–ice* and *–uce* both have “c” as the penultimate letter), and the items selected with those endings were the same length (e.g., *service* and *lettuce*). This ensured that target letters were always in the same position within words (Guérard, Saint-Aubin, Poirier, & Demetriou, 2012). In total, we selected 10 critical words (five per condition), each of which appeared several times within its passage. Each word had only one instance of the target letter. Log-transformed item frequency was balanced across conditions, $t(5.27) = 0.47$, $p = .96$ (Baayen et al., 1995).

Adults' texts. The three texts read by adults were written as nonfiction informational passages. One passage outlined the history of the Daredevil comic franchise (target letter: *u*; critical words: *figure*, *feature* [consistent]; *volume*, *costume* [inconsistent]). A second passage took the format of a news article, and described chef classes being offered by a restaurant (target letter: *c*; critical words: *service*, *haddock* [consistent]; *lettuce*, *product* [inconsistent]). A final passage summarized the process for obtaining a patent (target letter: *n*; critical words: *patent* [consistent], *legend* [inconsistent]). Critical words were repeated 3–4 times within passages, yielding a total of 16 items per condition. The order in which the passages were presented was fully counterbalanced across participants.

The passages were written such that critical words were always preceded by a function word (determiner, preposition, or particle). They were never preceded or followed by a word that contained the passage's target letter or by a word containing punctuation. Critical words never appeared at the beginning or end of a sentence.

The target letters also appeared in several non-critical words (see Table 5.1). We revised the texts to minimize these occurrences as much as possible; target letters appeared in less than 15% of words in a given passage. To ensure that the passages were readable and not overly contrived, we included a few instances of critical words that did not meet the criteria described above. Three appeared after content words ("action figure"; "black costume"; "poor man's patent"), and two were inflected forms of the word *patent* ("patented"; "patents"). These instances were not coded for target letter detection.

To ensure that our passages were written at an appropriate level of difficulty, we submitted them to TextEvaluator—a text analysis tool that provides automated passage complexity metrics (Napolitano, Sheehan, & Mundkowsky, 2015; Sheehan, Kostin, Napolitano, & Flor, 2014). The Overall Text Complexity scores for each passage are shown in Table 5.1. These scores fell within the ranges of complexity expected at the Grade 10–12 level, suggesting that their level of difficulty was appropriate for our undergraduate participants.

Table 5.1.

Summary of passage characteristics for the texts read by participants in Experiment 1 (adults) and Experiment 2 (children).

Passage (*target letter*)	# target letters: In critical words	# target letters: In non-critical words	Total # words	Text complexity score [a]
Adults' passages				
Daredevil comics (*u*)	14	51	711	990
Chef for a day (*c*)	12	53	726	1040
A guide to patents (*n*)	6	64	556	1100
Children's passages				
Trip to the dentist (*s*)	6	21	149	440
The talent show (*n*)	6	21	168	530
The birthday party (*u*)	6	9	144	520
Dinner at a restaurant (*c*)	6	12	182	580
Going for a walk (*u*)	6	17	170	480

[a] Text complexity scores were calculated using TextEvaluator (Napolitano et al., 2015). Scores for the adults' passages fell within the range of text complexity expected at a Grade 12 level (970–1360). Scores for the children's passages fell within the range of text complexity expected at a Grade 3 level (310–590).

Procedure. Adult participants were tested in a quiet university assessment room in groups of up to four people. Participants were seated to ensure that they could not see one another's work. They were instructed to read each passage silently for comprehension at their normal speed, and to concurrently search for the target letter specified at the top of each page. Participants were asked to mark each letter that they noticed while reading, and told that they could not go back to mark an undetected letter if they subsequently noticed it. After each passage, participants were asked to answer three multiple choice questions to assess their comprehension of the text. These instructions were provided in writing and explained verbally.

Throughout the session, participants were monitored by an experimenter who answered questions about the task and ensured that the participants were following instructions. In total, the testing session lasted approximately 30 minutes.

5.3.2 RESULTS

Data cleaning. Participants were scored on whether they detected the target letter in each critical word. As our outcome measure, we calculated participants' omission rates as the percentage of target letters that participants failed to detect. Participants' mean accuracy on the comprehension questions was 80.71% (*SD* = 14.96; 25% chance level).

Inspecting the letter omission rates separately for each text revealed a potential issue with the third passage (*'A guide to patents'*). Specifically, we saw a particularly low omission rate for items in that passage's consistent condition, in which participants were meant to detect the target letter *n* in the word *patent*. Examination of the descriptive statistics showed floor effects for Passage 3's consistent condition (*M* = 18.87%, *SD* = 28.87), and the omission rates for that condition also showed a clear positive skew (*z* = 4.71) far stronger than that seen in any other passage or condition. We suspect that this issue arose for two reasons: the critical word *patent* was the topic of the passage itself, and the word was repeated several times beyond the three instances that we coded (six times total). Both factors are likely to have made the word *patent* particularly salient to participants, increasing its letter detection rate. To corroborate our suspicions, we conducted secondary coding of all six instances of *patent*; doing so yielded the same floor effect (*M* = 22.01%, *SD* = 25.00) and pronounced skew (*z* = 4.50) as the original coding.

To address this issue, we based our analyses on target letter omission scores from the first two passages, for which there were no floor effects. The two remaining passages had a total of 13 items per condition; percent omission rates are shown in the right-hand panel of Figure 5.1. Importantly, after excluding this passage, our ending manipulation was maintained: associations with first-syllable stress were stronger for endings in the consistent condition ($M = .78$, $SD = .08$) than for endings in the inconsistent condition ($M = .33$, $SD = .06$), $t(6) = 9.08$, $p < .001$). Item frequency also remained balanced across conditions, $t(6) = 0.41$, $p = .70$.

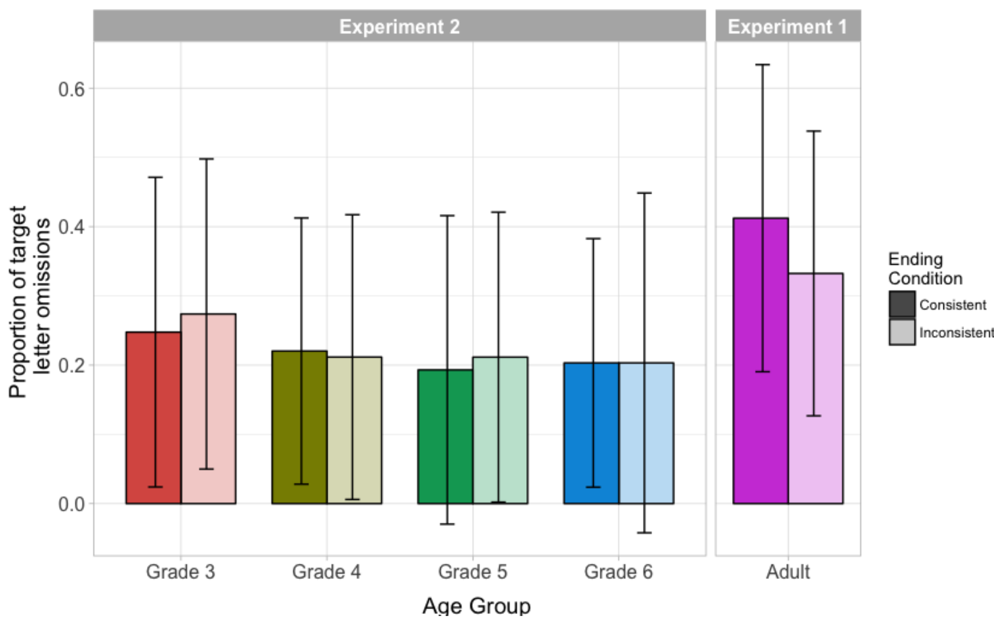


Figure 5.1. Mean rates at which participants omitted target letters as a function of ending condition. Error bars show standard deviations.

Examination of residual plots and *z*-scores showed that the condition totals had normal distributions with no outliers (skew and kurtosis were within an acceptable range of normality, per the guidelines of Tabachnick & Fidell, 2007). There were no missing data, as all participants identified the correct target letters throughout the task. Passage order had no effect on omission rates, nor did it interact with condition (*ps* > .42).

Analyses. To determine whether the association between word endings and lexical stress affects adults' processing of words in connected text, we conducted a paired samples *t*-test comparing participants' letter omission rates across conditions. This analysis showed a higher omission rate in the consistent condition (M = 41.22%, SD = 22.18) than the inconsistent condition (M = 33.24%, SD = 20.56), $t(52) = 2.97$, $p = .004$, $d = 0.41$. Figure 5.2 illustrates this effect, showing paired differences in participants' target letter omission rates across the consistent and inconsistent conditions. Supporting the by-participants analysis, we also found a nonsignificant but marginal trend when analyzing the data by items. An independent samples *t*-test showed a slightly higher omission rate for the consistent condition (M = 41.22%, SD = 9.55) than the inconsistent condition (M = 33.24%, SD = 13.11), $t(24) = 1.78$, $p = .089$, $d = 0.70$. To facilitate comparisons of our ending consistency effects with those reported in published studies of adults (e.g., Kelly et al., 1998; Mundy & Carroll, 2013), we also report eta-squared effect sizes. Eta-squared values were 0.145 and 0.117 for the analyses by participants and by items, respectively.

We present these findings with the caveat that they are based on a subset of the passages originally created for this experiment—a decision made based on our post hoc examination of the data. We argue above that removing a passage was appropriate, given the apparent issues with the critical word *patent*. However, we must bear this decision in mind when considering the results, particularly since the decision to remove the passage affected the pattern that we found. Repeating the analyses above with the full set of 16 items showed no effect of condition on letter detection, either by participant, $t(52) = 1.12$,

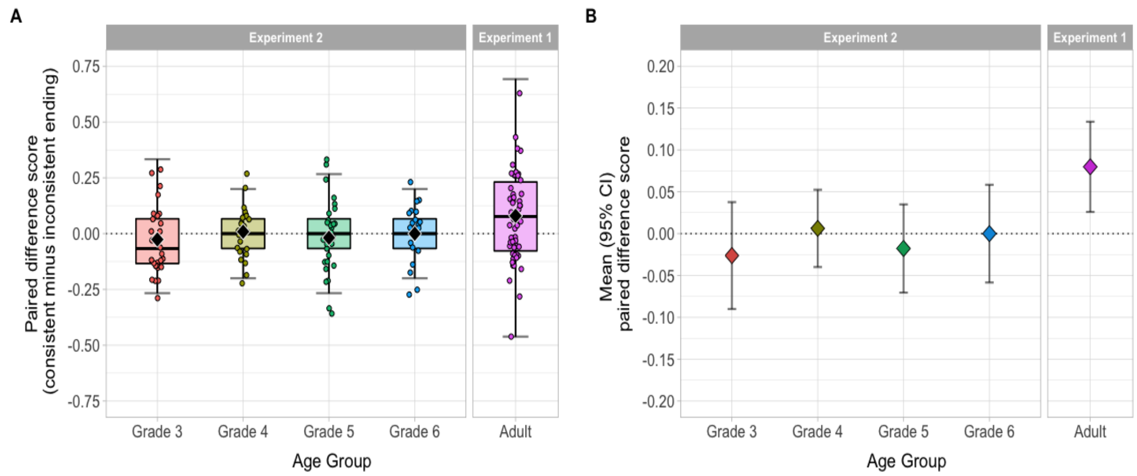


Figure 5.2. Paired difference scores in the rates at which participants failed to detect target letters in words whose endings were consistent versus inconsistent with their stress patterns. Panel A shows scatterplots and boxplots of paired differences in the rates at which each participant missed target letters in each condition. Panel B shows the corresponding mean differences for each age group; error bars represent 95% confidence intervals.

$p = .27$, $d = 0.15$ (consistent: $M = 37.03\%$, $SD = 20.51$; inconsistent: $M = 34.20\%$, $SD = 18.85\%$), or by item, $t(30) = 0.65$, $p = .52$, $d = 0.23$ (consistent: $M = 37.38\%$, $SD = 11.94$; inconsistent: $M = 34.55\%$, $SD = 12.14$).[17] Eta-squared values were 0.024 and 0.014 for the analyses by participants and by items, respectively. Given this analysis, we should interpret the main finding of this experiment with some caution.

Experiment 1 offers tentative evidence that adults' sensitivity to the association between word endings and lexical stress affects their processing of words in connected text. In Experiment 2, we examine whether the same is true of developing readers.

---

[17] Despite the skewness issues we noted above in Passage 3's consistent condition, the condition totals calculated across all three passages were normally distributed. Skew and kurtosis for these totals were within an acceptable range of normality, per the guidelines of Tabachnick & Fidell (2007).

5.4 EXPERIMENT 2: DEVELOPING READERS

5.4.1 METHOD

Participants. We recruited 106 elementary school students, Grades 3 to 6, from four schools in an urban region of eastern Canada. Data for five of these children were excluded from analyses, as they either ignored task instructions (n = 2) or did not complete the task (n = 3). Of the 101 included participants, 28 were in Grade 3 (mean age = 8.96 years; SD = 0.29), 23 were in Grade 4 (mean age = 9.91 years; SD = 0.35), 29 were in Grade 5 (mean age = 11.12 years; SD = 0.35), and 21 were in Grade 6 (mean age = 11.96 years; SD = 0.35). Testing took place in the spring, during the final months of the school year. Children received a small token of appreciation (e.g., a pencil or eraser) as thanks for their participation.

Materials. Five short texts were constructed for children, with an average length of 162 words (see Table 5.1). Passages were assigned a target letter, indicated to participants at the top of every page. The target letters were embedded in a set of critical words that we used to test for a missing letter effect.

Item selection. Item selection criteria were the same for Experiment 2 as they were for Experiment 1. Critical words were disyllabic nouns pronounced with first-syllable stress; half had endings associated with first-syllable stress (consistent condition), and half had endings associated with second-syllable stress (inconsistent condition). Associations between endings and stress were calculated from a corpus of elementary grade-level texts (Zeno, Ivens, Millard, & Duvvuri, 1995). Confirming our manipulation, the association with first-syllable stress was significantly higher for endings in the consistent condition (M = .75, SD = 0.02) than for endings in the inconsistent condition (*M* = .38, *SD* = 0.05), $t(8) = 14.44$, $p < .001$.

In total, we selected 10 critical words (five per condition). Each word had only one instance of the target letter. Item frequency was balanced across conditions at each grade, *t*s < 0.48, *p*s > .65 (Zeno et al., 1995).

Children's texts. The five passages read by children were written as journal entries in which a boy describes the events of his day. In the passages, he visits the dentist (target letter: *s*; critical words: *dentist*; *purpose*)[18], watches a talent show (target letter: *n*; critical words: *talent*; *legend*), attends a birthday party (target letter: *u*; critical words: *feature*; *costume*), eats at a restaurant (target letter: *c*; critical words: *service*; *lettuce*), and goes for a walk (target letter: *u*; critical words: *nature*; *volume*). Each critical word was repeated three times in its passage, for a total of 15 items per condition. The order in which passages were presented was counterbalanced across participants using a balanced Latin Square design.

As in Experiment 1, the texts were written such that the critical words were always preceded by a function word (determiner, preposition, or particle) and were never preceded or followed by a word that contained the passage's target letter. Critical words never appeared at the beginning or end of a sentence or next to punctuation. We minimized the number of target letters appearing in non-critical words to less than 15% of words in a passage, as shown in Table 5.1. All instances of Experiment 2's critical words were coded for letter detection, as none of the passages had additional repetitions of the words.

To ensure that our passages were written at an appropriate level for our youngest

[18] Critical words for each passage are shown with the consistent condition's item first and the inconsistent condition's item second.

participants, we obtained text complexity scores from TextEvaluator (see Table 5.1). All passages' scores fell within the range of complexity expected at a Grade 3 level. They also fell within ranges expected in Grades 4 and 5, and slightly below those of Grade 6, suggesting that the texts were appropriate for our older participants as well (Napolitano et al., 2015; Sheehan et al., 2014).

Procedure. Child participants were tested in groups at their schools, either in their classrooms or in a quiet assessment room. Participants were seated to ensure that they could not see one another's work. They were instructed to read each passage silently for comprehension at their normal speed, while also searching for the target letter specified at the top of each page. Participants were asked to circle each letter that they detected, and told that they could not go back to mark an undetected letter if they subsequently noticed it. After each passage, participants were asked to answer two multiple choice questions about what they had read. These instructions were provided in writing and explained verbally. To ensure that children fully understood the task, an experimenter demonstrated a practice sentence on a poster board.

Throughout the session, participants were monitored by two experimenters who answered participants' questions about the task and ensured that they were following instructions. In total, the testing session lasted approximately 25 minutes.

5.4.2 RESULTS

Data cleaning. As in Experiment 1, participants were scored on whether they detected the target letter in each critical word, and our outcome measure was their percent omission rate. One participant circled the wrong letter in a passage, but otherwise completed the task correctly; we treated their data for that passage as missing and imputed the participant's mean scores by condition. The left-hand panel of Figure 5.1 shows percent target letter omission rates by grade. Across grade levels, participants' mean accuracy on the comprehension questions ranged from 87.86% (SD = 13.15) to 98.28% (SD = 4.68; 33% chance level).

Two of the texts read by children have the same conceptual issue as the *patent* passage described in Experiment 1, in that one of their critical words was the topic of the passage (*dentist*; *talent*). Given that, we inspected omission rates by passage and grade. Floor effects in letter omission rates were widespread, appearing across grades in four of the five passages (the *feature–costume* passage being the only exception). While this characteristic of our data is noteworthy, its presence across the texts did not give us reason to exclude specific passages. As such, we report analyses that are based on all 15 items presented to children.

Across all passages, examining residual plots and z-scores of children's total omission rates by grade showed some instances of positive skew (consistent condition: Grade 3 $z = 2.99$; Grade 5 $z = 3.18$; inconsistent condition: Grade 5 $z = 3.24$; Grade 6 $z = 3.63$; these values deviate significantly from normality at $p = .01$ per Tabachnick & Fidell, 2007). We nevertheless conducted our primary analysis on these total scores, as ANOVA is reasonably robust to slight deviations from normality (Harwell, Rubinstein, Hayes, & Olds, 1992; Schmider, Ziegler, Danay, Beyer, & Bühner, 2010). However, to corroborate our findings, we repeated the analysis on square-root transformed omission rates. The transformations corrected all skewness issues, and analyzing the transformed scores did not change the pattern of results reported below. Passage order had no effect on omission rates, nor did order interact with condition (ps > .15).

Analyses. To determine whether the association between word endings and lexical stress affects children's processing of words in connected text, we conducted a 4 x 2 ANOVA with grade level (3; 4; 5; 6) as a between-subjects factor and condition (consistent; inconsistent) as a within-subjects factor. The main effect of condition was nonsignificant, $F(1, 97) = 0.46$, $p = .50$, $\eta_p^2 = 0.01$, as was the main effect of grade, $F(3, 97) = 0.52$, $p = .67$, $\eta_p^2 = 0.02$. Likewise, the interaction between condition and grade was nonsignificant, $F(3, 97) = 0.30$, $p = .83$, $\eta_p^2 = 0.01$.[19] The paired difference scores presented in Figure 5.2 illustrate these null effects. A by-items analysis similarly showed only null effects (*p*s > .55).

5.5 GENERAL DISCUSSION

Lexical stress is a key feature of the phonology and lexical identity of multisyllabic words in English, but the role it plays in reading has only recently been given attention. Much of the work to date has focused on the written cues that readers use to read or recognize isolated words (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). For example, there is evidence that readers draw on the associations between word endings and lexical stress, responding to words whose endings are consistent with their stress patterns more quickly and accurately than to words whose endings are not (Kelly et al., 1998; Mundy & Carroll, 2013). It has been suggested that these effects emerge because word endings' links with stress affect lexical access: inconsistent endings offer misleading cues to stress, impeding lexical access, whereas consistent endings correctly cue stress, facilitating lexical access (Kelly et al., 1998). However, the extent to

[19] When analyses were repeated using square-root transformed omission rates, we found the same pattern of nonsignificant results. Condition main effect: $F(1, 97) = 0.17$, $p = .69$, $\eta_p^2 = 0.002$. Grade main effect: $F(3, 97) = 0.62$, $p = .60$, $\eta_p^2 = 0.02$. Interaction: $F(3, 97) = 0.94$, $p = .42$, $\eta_p^2 = 0.028$.

which readers draw on these cues while reading for comprehension has not previously been tested, leaving open questions about the scope of ending consistency effects on reading in English. The current study addressed this gap, using a letter-detection paradigm to determine whether endings act as cues to lexical stress when words are read in connected text. We found affirmative evidence for this across two of the passages read by adult readers, who detected fewer target letters in words whose endings were consistent with their stress pattern than in words whose endings were inconsistent. This pattern suggests that endings that are consistent with stress facilitate more efficient lexical access than endings that are not, yielding a higher letter omission rate in adults (Klein & Saint-Aubin, 2016). On the other hand, word endings' consistency with stress had no apparent effect on the letter detection rates of children in Grades 3 to 6.

5.5.1 ENDING SPELLINGS AS STRESS CUES IN CONNECTED TEXTS

In Experiment 1, word endings' consistency with stress affected the rate at which adults detect target letters while reading connected text for comprehension (specifically, this finding emerged in our primary analysis of the passages that did not have floor effects). That is, in the analysis by participants, adults identified fewer target letters in words whose endings were consistent with their stress pattern than in words whose endings were inconsistent with their stress pattern. When interpreted in light of the view that higher letter omission rates occur when lexical access is quick (Klein & Saint-Aubin, 2016), this finding suggests that word endings can indeed affect readers' processing of lexical stress when reading words in connected text. And notably, the pattern of results that we found in Experiment 1 aligns with the existing research on adults' processing of stress in isolated words. Just as the adult participants in our study tended to detect fewer target letters when words' endings were consistent with their stress pattern, adults in prior work made quicker and more accurate naming and lexical decision responses toward words with consistent endings (Kelly et al., 1998; Mundy & Carroll, 2013).

On the other hand, our findings in Experiment 2 diverge somewhat from the prior literature on children's use of word endings to guide stress placement (e.g., Arciuli et al., 2010). In our data, it was clear that the word ending manipulation had no effect on children's letter detection at any of the grade levels we examined. Although we must be cautious of interpreting null findings, the effect sizes in Experiment 2 were quite small (η_p^2 = .01 to .02). This supports the conclusion that children did not draw on word endings to process stress while reading words in connected text—a finding that emerges in spite of the evidence that children are sensitive to the associations between ending spellings and lexical stress in English from a relatively young age. Specifically, prior work by Arciuli and colleagues (2010) shows that word endings affect how children place stress in unfamiliar pseudowords. Indeed, in that study, the effect of word endings on stress placement was particularly strong among 9- to 12-year-old children—the same approximate age range that we included here (our child participants were, on average, 8.9 to 11.9 years of age). We note this discrepancy with the caveat that the body of research done with English-speaking children is quite small, and our method is certainly quite different from the pseudoword naming approach used by Arciuli and colleagues (2010). Further research will thus help to reconcile the discrepant findings. Even so, our results suggest elementary school-aged children do not use word endings to process stress when they read in connected text.

Another useful way to place our study's findings within the literature on word endings as cues to stress is to compare the magnitude of effects seen in the current study with those found in prior work. In doing so, we focus on the existing research in adults, as those are the studies for which the information necessary to calculate effect size estimates is available (Arciuli & Cupples, 2006; Kelly et al., 1998; Mundy & Carroll, 2013). For reference, we have provided effect size estimates and confidence intervals for each of these studies in Appendix G, along with the corresponding effect size from Experiment 1. As the appendix shows, past work has reported relatively large effects of ending cue consistency on adults' accuracy in naming and lexical decision tasks (e.g., η_p^2 point estimates of approximately .70 for real words and .58 for nonwords; Arciuli & Cupples, 2006; Kelly et al., 1998; Mundy & Carroll, 2013). These effects are noticeably larger than the effect we found in Experiment 1 ($\eta_p^2 = .145$)—a fact that becomes particularly clear when we compare the effect size confidence intervals of Experiment 1 and measures of accuracy in isolated words. Indeed, for the most part, the letter detection and accuracy confidence intervals do not overlap.[20] This discrepancy is particularly noteworthy from a theoretical standpoint, as stress placement accuracy is the primary focus of the existing models of multisyllabic word reading (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009).

In contrast to the comparisons with accuracy, Experiment 1's effect size confidence interval does overlap with the intervals capturing word endings' effect on response times to isolated words. This suggests that the impact of word endings on our

[20] The one exception to this statement is Kelly et al. (1998)'s Experiment 3. In that experiment, items were higher frequency than in Experiments 1 and 2 of the study. Participants' accuracy rates were quite high, and response times were shorter than in the study's other experiments.

letter-detection task more closely resembles the speed at which readers recognize and name isolated words (which is itself attenuated by cognitive noise; Diependaele, Brysbaert, & Neri, 2012) than the accuracy with which they place stress in those words. Yet even so, it is worth noting that Experiment 1's effect size was numerically smaller than the point estimates found for reaction times to isolated words (η_p^2 estimates of approximately .26).

Taken together, these comparisons suggest that endings' consistency with stress has a smaller effect on adults' processing of words in connected text than on their processing of isolated words. When noting this discrepancy, we should be mindful of the fact that various factors can affect the magnitude and precision of effect size estimates (e.g., Brand & Bradley, 2016), including the number of participants and/or items involved in a study. In the current study, methodological constraints limited us to a relatively small number of target items. In contrast, prior naming and lexical decision studies have typically been able to include far more (Kelly et al., 1998; Mundy & Carroll, 2013). As such, it is possible that Experiment 1's relatively small effect size is a simple artifact of our study design. Future research that uses a range of experimental paradigms will be needed to establish the true magnitude of the effect that word endings' consistency with stress has on reading in connected text—we discuss this point in more detail below. However, in the absence of this additional research, it is worth exploring conceptual explanations for the discrepant effect sizes across studies of isolated word reading and reading in text.

The possibility that reading words in context might attenuate the role of ending spellings as cues to stress has not been well explored to date. However, this idea is consistent with one recent study, which was conducted with Italian-speaking adults (Spinelli, Sulpizio, Primativo, & Burani, 2016). In Italian, it is possible to establish contrasts between the stress pattern suggested by an ending's spelling and the stress pattern suggested by a specific morpho-syntactic use of that ending. For example, most Italian words that end in *–ita* have penultimate stress (e.g., *saLIta*). Among those words, however, are a subset of cases where *–ita* serves as a grammatical marker for a third-person indicative verb; within that subset, most words have antepenultimate stress (e.g., *DUbita*). When Spinelli and colleagues presented nonwords in isolation, readers tended to pronounce them with the stress pattern most strongly associated with their ending spelling (e.g., penultimate stress for nonwords ending in *–ita*; see also Burani & Arduino, 2004; Burani et al., 2014; Colombo et al., 2014; Sulpizio, Arduino, Paizi, & Burani, 2013). Interestingly, though, when the same nonwords were shown after a context word (e.g., a third-person pronoun), readers increasingly used the stress pattern associated with the grammatical properties of the ending (e.g., antepenultimate stress for nonwords ending in *–ita*). Spinelli and colleagues' (2016) findings do not address our study's goal of looking at endings as stress cues in a naturalistic reading context—they asked participants to read nonwords, and the only step toward providing connected text was a single context word. Even so, their findings suggest that elements of sentence context (which are available in connected text) can affect the extent to which readers use ending spellings as cues to stress placement.

If we are right in suggesting that word endings play a smaller role as stress cues when reading in connected text than in isolated words, then it is important to consider why that might be the case. One plausible explanation is that meaningful texts offer a wider variety of stress cues than are available for isolated words, leading readers to draw less heavily on written cues like ending spellings. To an extent, we saw this at play in Spinelli and colleagues' (2016) study, where clear syntactic markers served as stress cues that contrasted with—and attenuated the effect of—their word endings' broad associations with lexical stress. Although the particular stress information used in that study was specific to Italian, English also has a variety of different probabilistic cues to stress. For instance, stress in English is associated with grammatical category, such that disyllabic nouns tend to have first-syllable stress while verbs tend to have second-syllable stress (Kelly & Bock, 1988). When reading texts, syntactic information can provide reliable cues to a word's grammatical category, and thus its likely stress pattern (Breen & Clifton, 2011). Likewise, when a word itself can be accurately predicted from context or quickly accessed from the lexicon, knowledge of its specific stress pattern may override other, more probabilistic, cues (Harris & Perfetti, 2016; Perry et al., 2010). A combination of these factors, among others, might serve to make word endings less important as stress cues, broadly reducing a reader's need to rely on them, though of course confirming this speculation will require further research. In any case, our study's salient finding is that the prior work on isolated word reading likely overestimates the effect of endings as stress cues during the everyday task of reading for comprehension.

5.5.2 EXTENDING MODELS OF ISOLATED MULTISYLLABIC WORD READING

The current study helps to clarify the scope of endings' role as stress cues in English, beyond what is captured by the existing models of multisyllabic word reading that have inspired much of the recent empirical work on the subject. For the most part, these models focus on accurate stress placement as their output,[21] and they are reasonably successful at using orthographic cues to mirror human performance (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). There has been a fair bit of empirical research aiming to test and build upon these models, in which the focus understandably remains on stress placement in isolated words and pseudowords (e.g., Arciuli et al., 2010; Mousikou, Sadat, Lucas, & Rastle, 2017). Yet because the existing models of multisyllabic word reading deal with isolated words, they assign stress based solely on combinations of orthographic, phonemic, and/or lexically-derived stress cues that are present in a word. These models cannot include any of the attenuating factors that we argue are present in connected text, and thus may inflate the role of ending cues in everyday reading for comprehension. Indeed, as shown in the appendix, adult readers' accuracy outcomes—central to all existing models—have noticeably larger ending consistency effects than those found in the current study. Thus, in order to arrive at models that more fully capture the realities of everyday reading for comprehension, we see value in expanding the scope of empirical work to explore the process of reading words in connected text. Doing so was a key contribution of the current study. By continuing this line of research, we will be better able to characterize—and ultimately model—the role of endings as stress cues across a variety of reading tasks.

5.5.3 LIMITATIONS AND FUTURE DIRECTIONS

When interpreting our adult findings, we need to remain aware that word endings' consistency with stress had no effect on letter detection rates when we analyzed all three

[21] Only one existing model attempts to capture reaction time estimates (CDP++; Perry et al., 2010), which is the outcome measure for which the effects of ending consistency are somewhat closer to our study's. Even in that case, however, the model's reaction times are only modestly associated those of adult readers (Mousikou et al., 2017).

of Experiment 1's passages. As argued earlier, we believe that this was due to the floor effects in one passage; accordingly, we removed that passage from our primary analysis. Although we believe that our decision to remove the passage with floor effects was appropriate, further replication of this work will help to increase confidence in the results. With that caveat in mind, our interpretation of Experiment 1's finding suggests that the effect of ending cue consistency when reading for comprehension (as reported in the current study) is somewhat attenuated when compared with that of reading words in isolation (as reported in prior work; e.g., Kelly et al., 1998; Mundy & Carroll, 2013). In that respect, our conclusions are similar regardless of which analysis one chooses to interpret. Either the effect of endings' consistency with stress is modest but significant (when excluding the problematic third passage), or it is more strongly attenuated, to the point of being nonsignificant (when including all passages). Both interpretations suggest that reading text for comprehension reduces adults' use of word endings as cues to lexical stress. The current study is a valuable starting point for research that focuses on reading in connected texts, but this line of research will certainly benefit from further work to clarify our findings and allow for more concrete interpretations.

Our primary goal for this study was to explore whether word endings act as cues to lexical stress when reading connected text for comprehension. We did so by placing carefully-chosen target words in realistic, age-appropriate connected texts, which allowed us to explore our research question in a naturalistic reading context. That said, a useful extension of the current study would be to explore a range of text manipulations that might affect readers' use of word endings as cues to stress. Word predictability is one possible manipulation. Prior research has established that readers use predictable sentence contexts to form expectations about the stress patterns of upcoming words (e.g., the word *system* in a sentence beginning *'He was enraptured by news of the space program and the solar ____'*; Harris & Perfetti, 2016). When a word's identity is highly predictable, its stress pattern can be quickly and accurately retrieved from the lexicon, reducing the need for readers to rely on sublexical stress cues (as suggested by Perry et al., 2010). Drawing on this, we might expect to see a relatively large effect of endings' stress consistency when items are preceded by low-predictability words, and a smaller or nonsignificant effect of endings in items preceded by high-predictability words (e.g., *solar* + *system*). Testing this prediction would be a valuable step toward explaining why the role of word endings as cues to stress is smaller when reading words in text than when reading them in isolation.

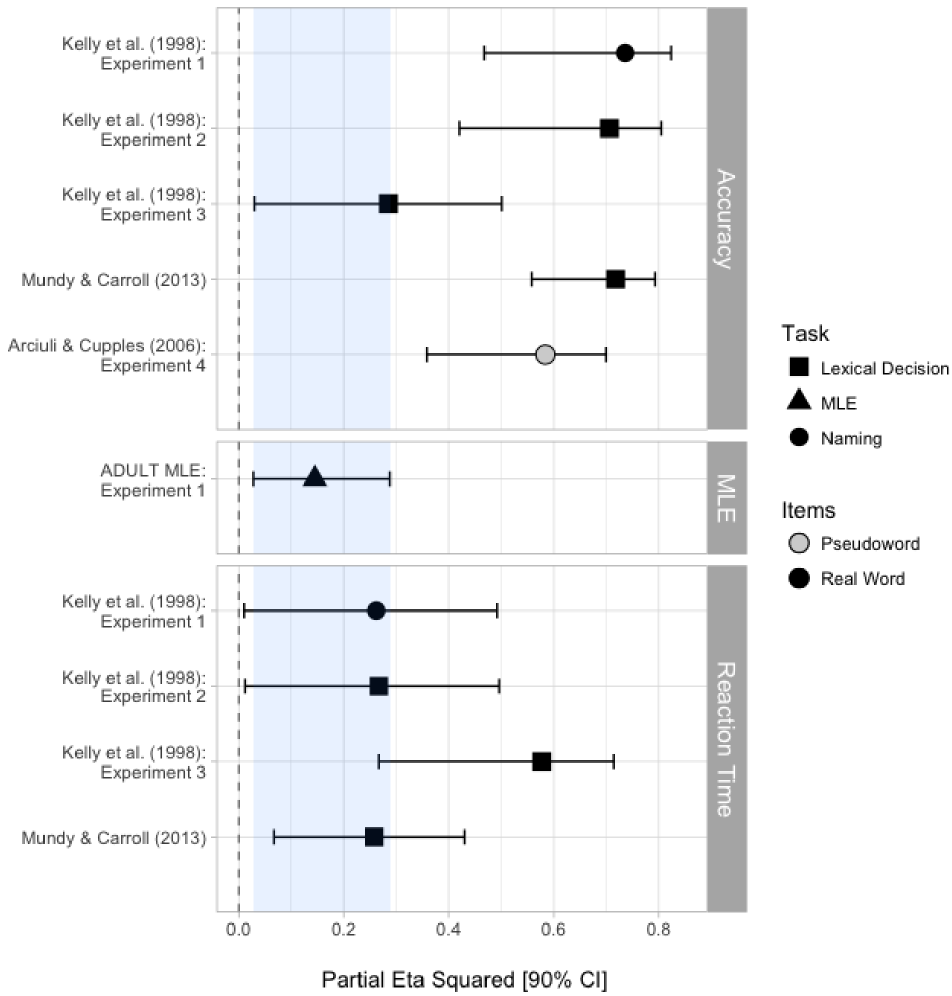
The letter-detection paradigm that we used in the current study was valuable, as it gave us a simple and effective means by which to explore readers' use of stress cues in connected text. However, future research that makes use of other online measures of text reading will help to address some of our task's limitations. A first limitation is that testing for a missing letter effect put stringent constraints on item selection. This level of control was important to our study's design, but it dramatically reduced the number of viable items available to us. For example, not only did items need to be matched on word frequency, length, and ending cue strength (as would be true in any paradigm), they also could only contain one instance of the target letter, which had to occur in the same position across words (a constraint specific to letter-detection tasks). This may have contributed to the marginal item effect that we saw in Experiment 1. Using other methodologies in which these constraints do not apply would allow us to work with a larger and more diverse set of stimuli, helping to ensure the generalizability of results. Secondly, using different paradigms would allow for direct manipulation of text characteristics (e.g., word predictability) that were not possible in the current study. We were not able to include such text manipulations here, because the best practices of the missing letter effect paradigm require items to be preceded by function words (e.g., *a*; *the*) that provide very little information about the identity of the word that follows. Exploring these questions in with other paradigms will therefore help to test some of the speculation we provided when interpreting the magnitude of findings in our study. Fortunately, there are several viable ways to measure how readers process stress during silent reading—for instance, researchers have done so using eye-tracking (Ashby & Clifton, 2005) and ERP paradigms (Kriukova & Mani, 2016). To date, though, these methods have not been applied to questions of how written stress cues affect the processing of stress in connected text. Doing so would provide valuable evidence to support or extend our study's findings.

In sum, the current study helps to clarify the role of word endings as cues to lexical stress in multisyllabic written words. It builds on prior evidence that adult readers are sensitive to the links between ending spellings and lexical stress in English (Arciuli & Cupples, 2006; Kelly et al., 1998; Mundy & Carroll, 2013), demonstrating that this sensitivity plays a small but significant role when adults read connected texts for comprehension. This provides a crucial look beyond the domain of isolated word reading, which has been the emphasis of the literature to date. Yet relative to that prior literature, the magnitude of our findings suggests that reading connected text might attenuate adults' use of word endings as cues to stress. In a similar vein, our second experiment found that word endings' consistency with stress did not affect children's processing of words in connected text, despite prior evidence that children use word endings to place stress in isolated pseudowords (Arciuli et al., 2010). Together, the discrepancies between our findings and those of prior studies provide a caution against generalizing from research that is limited to isolated word recognition, as that work seems to overestimate the role that endings play in reading connected text. As such, our findings highlight the value of exploring the processes involved in multisyllabic word reading using tasks that approximate everyday reading for comprehension.

5.6 APPENDIX G: SUMMARY OF WORD ENDINGS' STRESS CONSISTENCY EFFECT SIZES

Figure 5.3. Forest plot showing the effects of ending cue consistency in published studies of English-speaking adults. Dots show the by-participant effect size point estimates (η_p^2), with error bars showing 90% confidence intervals for each effect size.

Effects were derived from prior studies on isolated words (lexical decision; real word naming; pseudoword naming), and they are shown in comparison with the ending consistency effect seen in Experiment 1 of the current study (Adult MLE). To facilitate this comparison, the plot's shaded region shows the confidence interval for the Adult MLE study's effect size.



Notes. MLE = missing letter effect. Plot shows the effect in each study that best captures ending cue consistency: Kelly, Morris, & Verrekia (1998): stress x spelling interaction; Mundy & Carroll (2013): orthographic reliability main effect; Arciuli & Cupples (2006): ending type main effect. No other ending-related effects were significant in these studies.

CHAPTER 6
DISCUSSION

6.1 SUMMARY OF DISSERTATION GOALS

The overarching goal of this dissertation was to explore the role of written word endings as orthographic cues to stress in English words. Across four studies, I sought to clarify the nature and scope of word endings' role as cues to stress in disyllabic English words across a variety of reading tasks. Chapter 2 explored the claim that suffix-like word endings act as particularly strong cues to stress by comparing the extent to which suffix and non-suffix endings affect visual word recognition. Chapter 3 tested predictions derived from statistical learning, a mechanism by which readers might learn of the associations that exist between stress and spelling in English. Chapter 4 considered the specificity with which readers associate stress and spelling, establishing whether sensitivity to word endings as stress cues exists independently of vowel quality. Finally, Chapter 5 explored the scope of word endings' ability to cue stress in English, testing whether they do so when words are read in connected text.

Taken together, these studies inform our understanding of written cues to stress among adults (Chapters 2 through 5) and developing readers (Chapters 4 and 5). Beyond their novel empirical insights, these studies address key theoretical questions raised by competing models of disyllabic word reading in English (Perry, Ziegler, & Zorzi, 2010; Rastle & Coltheart, 2000; Ševa, Monaghan, & Arciuli, 2009). For reference, Table 1.1 provides an overview of each model's stance on this dissertation's overarching questions. In the sections that follow, I review the findings of each study and discuss their implications for theory.

6.2 THEORETICAL CONTRIBUTIONS

6.2.1 SUFFIX AND NON-SUFFIX ENDINGS AS STRESS CUES

The experiments in Chapter 2 tested a key point of contrast between the existing models of disyllabic word reading—namely, whether suffix-like word endings play a special role as cues to stress when compared with non-suffix endings. According to the rule-based view put forward in Rastle and Coltheart's (2000) dual-route model, suffix-like endings (which have the spelling of a meaningful English suffix; e.g., *–ar*) play a uniquely strong role as cues to stress in English. In fact, non-suffix endings (e.g., *–el*) are not included in their model, and thus cannot be used to guide stress placement in written words. Based on this account, we would expect suffix-like endings (e.g., *–ar*) to act as stronger cues to stress than non-suffix endings (e.g., *–el*). By contrast, the connectionist and CDP++ models both claim that suffixes play no special role in stress assignment (Perry et al., 2010; Ševa et al., 2009). If these two models are correct, we would expect all word endings—suffix-like or not—to have similar effects on readers' processing of stress in written words.

To test these contrasting predictions, adults completed visual lexical decision tasks in which some items had endings that accurately cued their stress patterns, whereas others had endings that inaccurately cued stress. In both experiments, we replicated the finding that English-speaking adults respond more quickly to words whose endings are consistent with their stress patterns than words whose endings are inconsistent with their stress patterns (e.g., Kelly, Morris, & Verrekia, 1998; Mundy & Carroll, 2013). More notably, though, we also showed that that the effect of an ending's consistency with stress exists across ending types. The key theoretical contribution of this chapter came from

Experiment 1, where pseudosuffix endings' consistency with stress (e.g., *–ar* in *MORTar* vs. *guiTAR*) had the same effect on readers' lexical decisions as did non-suffix endings' consistency with stress (e.g., *–el* in *NOVel* vs. *hoTEL*). This finding is in direct contrast with the underlying assumption of Rastle and Coltheart's (2000) model, which was built on the premise that suffix-like word endings are uniquely strong cues to stress in English. Instead, our results support the view that all word endings—regardless of whether they can be used as suffixes—act as cues to stress in English (Perry et al., 2010; Ševa et al., 2009).

Although practical constraints prevented us from testing for this interaction when using words with genuine suffixes, the use of pseudosuffix endings effectively tested the contrasting theoretical predictions, as Rastle and Coltheart's (2000) model did not differentiate between real and pseudosuffix uses of endings. In Experiment 2, however, we were able to show that readers processed suffix-like endings similarly, regardless of whether they appeared in words as real suffixes (e.g., *BEGGar*) or as pseudosuffixes (e.g., *MORTar*). This finding serves two functions: it supports Rastle and Coltheart's (2000) decision to treat the two ending types similarly in their model, and more broadly, it adds further evidence that readers process word endings similarly, whether they are suffixes, pseudosuffixes, or non-suffix endings (per Perry et al., 2010; Ševa et al., 2009).

Taken together, the findings presented in Chapter 2 provide support for the connectionist and CDP++ models' view of stress assignment in disyllabic written words over Rastle and Coltheart's rule-based model. However, it is important to be clear about what specific conclusions our findings support. When interpreting these findings empirically, I would certainly not dispute that suffixes play an important role in English stress assignment. As noted previously, a role of suffixes in determining stress placement has been established linguistically (Fudge, 1984) and behaviourally among both adults (e.g., Wade-Woolley & Heggie, 2015) and children (e.g., Jarmulowicz, 2006). With this in mind, I argue that Chapter 2's key empirical contribution is in showing that suffixes are not unique in their role as stress cues. Non-suffix endings with similarly strong associations with stress have just as strong an effect on readers' processing of multisyllabic words. And from a theoretical standpoint, my findings cannot be used to categorically support a connectionist mechanism for stress assignment over a rule-based one. In principle, it might be possible to develop a rule-based model that incorporates a wider range of ending cues (though efforts to do so would likely be complicated by the need to precisely define the ending units most relevant to stress assignment in English; see section 6.4.1 for more discussion of this point). Instead, what my findings refute is a purely suffix-based—or even pseudo-suffix based—set of rules for stress assignment.

6.2.2 MECHANISM BEHIND READERS' LEARNING OF STRESS CUES

Chapter 3's findings take these conclusions a step further by providing more general support for a probabilistic, statistically-based view of endings' role as stress cues (Perry et al., 2010; Ševa et al., 2009) over a deterministic, rule-based one (Rastle & Coltheart, 2000). To recap, Ševa and colleagues' connectionist model and the sublexical route of CDP++ each propose that statistical learning drives readers' sensitivity to word endings as stress cues. Through exposure to a variety of words with a given ending cue, the models extract broader regularities about stress assignment that probabilistically weigh their encounters with both positive examples of a regularity (e.g., *NOVel* supports the link between *–el* and first-syllable stress) and negative counterexamples to that regularity (e.g., *hoTEL*, which contradicts that overall pattern). Under this theoretical perspective, we would expect readers to respond more quickly to words that have a large number of stress friends (e.g., for the word *NOVel*: *ANGel* and *BARRel*). At the same time, we would expect responses to be somewhat slowed when words have a large number of stress enemies (e.g., *hoTEL*).

In contrast, Rastle and Coltheart's (2000) rule-based model treats stress cues as deterministic; word endings in the model's affix store either take stress or they do not. This classification is based on the stress pattern most strongly associated with a given ending, and it overlooks the graded nature of those associations by creating a binary stress assignment rule for each ending. Rastle and Coltheart's model did not explicitly propose a mechanism by which readers come to learn of rules to stress assignment. However, their approach suggests that once these rules have been acquired, readers will not continue to draw on their respective levels of exposure to stress friends and stress enemies when processing stress in written words. That is, we would not expect both number of stress friends and number of stress enemies to independently affect readers' lexical decisions.

Our second research question in Chapter 3 explored these differing predictions, and showed support for the connectionist and CDP++ models' probabilistic approach (Perry et al., 2010; Ševa et al., 2009). That is, after accounting for a word's number of stress friends, that word's number of stress enemies contributed unique variance to readers' lexical decision latencies. Likewise, after accounting for number of stress enemies, a word's number of stress friends predicted unique variance in response times.

The two variables' effects were of similar size, but were in opposing directions: a large number of stress friends facilitated quick responses, whereas a large number of stress enemies slowed responses. This pattern of results fits nicely with the claim that statistical learning is a mechanism behind readers' use of word endings as stress cues in English. Specifically, it shows that readers draw on their exposure to all instances of a given ending when processing stress in written words.

With that being said, we did not find evidence for a stronger prediction derived from the connectionist approach: that readers would be particularly sensitive to the ending cues that they encounter most often. This prediction was addressed by the analyses that tested the respective roles of a word's proportion and number of stress friends on readers' lexical decision latencies. As expected, the higher an item's proportion of stress friends (i.e., the stronger an ending's association with lexical stress), the quicker participants were to respond. However, after accounting for proportion of stress friends, a word's number of stress friends had no effect on lexical decisions. That is, words with commonly-occurring ending cues (e.g., *NOVel*) had the same impact as did words with comparatively rare ending cues (e.g., *ballOON*). This is not what we would have expected based on the view that readers' sensitivity to the patterns in written English develops through accumulated exposure to the various regularities in text. As such, Chapter 3's findings are perhaps only qualified support for predictions of the connectionist models and the statistical learning mechanism that underlies them—at least among skilled adult readers. Future work that explores this issue among developing readers may help us to further test the mechanistic predictions of the connectionist and CDP++ models among readers with less overall exposure to English text.

6.2.3 VOWEL QUALITY AND STRESS PLACEMENT

In addition to testing claims of the existing models of disyllabic word reading, the findings presented in this dissertation explore gaps in those models. The extent to which the effects of stress and vowel quality can be separated when processing written words is one such gap. On this issue, the three models each take a different approach. For instance, Rastle and Coltheart's (2000) dual-route model concurrently assigns stress and determines vowel quality in the written words it encounters. This approach models the tight link between stress and vowel quality in English, but makes it difficult to isolate the role of stress itself when reading multisyllabic words. In contrast, CDP++ represents stress and vowels independently of each other—the model's sublexical route learns of the links between stress and spelling in a way that mirrors, but is fully independent of, its learning of segmental grapheme–phoneme correspondences (Perry et al., 2010). This approach isolates lexical stress, but has been difficult to verify empirically given how closely stress and vowels are linked in English. Finally, Ševa and colleagues' (2009) connectionist model focuses solely on lexical stress; it does not include segmental phonology in its training or output. As such, it cannot explore whether stress and vowel quality play separable roles when readers process written words. The wide range of strategies used by these models makes it clear that there is no theoretical consensus around how stress interacts with segmental phonology during word reading in English. It also highlights an important empirical question: namely, whether readers make use of the links between stress and spelling in English when vowel quality is removed from the equation.

In Chapter 4, I addressed this empirical question with a pseudoword spelling choice task that manipulated aural stress patterns while maintaining vowel quality across pronunciations. Although children in the early- to mid-elementary grades did not use stress to guide their spelling decisions, our older participants did. That is, adults and children in the late-elementary grades preferentially chose the ending spelling that was associated with the stress pattern of the pseudoword that was aurally presented to them (e.g., choosing *–et* after a pseudoword pronounced with first-syllable stress; choosing *–ette* after one pronounced with second-syllable stress). This provides compelling support for the claim that readers' use of the links between stress and spelling English is indeed, at least partially, attributable to stress—it is not solely an artifact of segmental vowel quality.

These findings help to validate the architecture of CDP++'s sublexical route (Perry et al., 2010). As noted, CDP++ learns of the associations between stress and spelling while separately learning grapheme–phoneme correspondences (which include the links between vowels and spelling). Crucially, this approach allows lexical stress to play a specific role in processing written words (e.g., Arciuli & Cupples, 2006; Arciuli, Monaghan, & Ševa, 2010; Kelly et al., 1998; Mundy & Carroll, 2013) without disputing the important role of vowel quality in recognizing and producing English words (e.g., Cutler, 1986; Cutler, 2015; Wood, 2006). As an encouraging bonus, this separability of stress and vowel quality suggests that models of English word reading might be structurally compatible with languages in which lexical stress is entirely separate from segmental phonology (e.g., Greek; Italian). In those languages, vowels are not reduced in unstressed syllables, and as such, we would certainly expect the effects of stress on processing written words to be distinct from those of vowel quality (see, for example, the implementation of CDP++ in Italian; Perry, Ziegler, & Zorzi, 2014). The results that I reported in Chapter 4 suggest that similar distinct processes could apply in English, despite the cross-linguistic differences in how lexical stress is realized (Cutler, 2015).

6.2.4 STRESS CUES IN CONNECTED TEXT

A final theoretical gap explored in this dissertation concerns the scope of endings' role as cues to stress in written English. More specifically, I sought to establish whether word endings act as stress cues when words are being read in connected texts. This question represents an important extension of the existing models of disyllabic word reading, all of which focus on stress placement when reading isolated words (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). The empirical work showing that word endings affect readers' processing of written words has likewise focused on words read in isolation (e.g., Arciuli et al., 2010; Kelly et al., 1998; Sulpizio & Colombo, 2013; Sulpizio, Arduino, Paizi, & Burani, 2013). Certainly, it is important that we understand the processes involved in isolated word reading and recognition. However, isolated word reading is quite removed from the everyday task of reading connected texts for comprehension (e.g., Kuperman, Drieghe, Keuleers, & Brysbaert, 2013). As such, it is theoretically and empirically important to determine whether word endings act as cues to stress when they are read in connected texts.

In Chapter 5, I addressed this question with a paradigm in which participants read texts for comprehension while also circling a specified target letter each time they noticed it. The patterns in readers' letter detection rates speak to the processes involved with reading words in connected text—when lexical access of words is relatively quick, we expect readers to detect fewer target letters than when lexical access is slowed (e.g., Klein & Saint-Aubin, 2016). With this in mind, the experiments in Chapter 5 drew on the claim that lexical access is facilitated when a word's ending is consistent with its stress pattern and is impeded when a word's ending is inconsistent with its stress pattern (Kelly et al., 1998). To date, this claim had only been explored during isolated word reading. In Chapter 5, I extended it to the reading of connected text by comparing letter detection rates in words whose endings were either consistent or inconsistent with their stress patterns. This manipulation had no effect on letter detection rates among children in the mid- to late-elementary grades. However, we found tentative evidence that adult readers do use word endings as cues to stress when reading words in connected text. Specifically, adults detected fewer target letters in words whose endings were consistent with their stress patterns than words whose endings were inconsistent with stress.

This finding in adults is beyond the purview of the existing models of disyllabic word reading—none were designed to extend beyond the reading of isolated words (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). This is not a limitation of the models, per se; their focus on isolated word reading is appropriate, particularly as an initial step toward understanding the cues to stress that exist in written English words. However, it is worth considering that the effect of endings' consistency on letter detection rates was fairly small in Chapter 5, particularly when compared with prior work into endings' effects on accuracy when reading isolated words (Kelly et al., 1998; Mundy & Carroll, 2013). This discrepancy in effect sizes is theoretically important, as all three of the existing models of disyllabic word reading focus heavily on stress placement accuracy when reading isolated words (Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). This suggests that the existing models—and the empirical studies that they inspire—might overestimate the role of word endings as cues to stress when reading connected texts. Future work that directly compares readers' use of word endings when reading words in isolation and in texts will help to define the scope of models of disyllabic word reading and will further our understanding of the processes involved in multisyllabic word reading during everyday reading tasks.

6.2.5 COMPARING MODELS ON DISSERTATION FINDINGS

Taken together, the findings presented in this dissertation inform the existing models of disyllabic word reading in novel ways. They establish that an effective theoretical account of English word reading must be able to: (1) recognize and use both suffix and non-suffix word endings as stress cues; (2) gradually learn of the associations between word endings and lexical stress through exposure to a variety of words that demonstrate those relationships (weighing both stress friends and stress enemies); (3) process the associations between word endings and lexical stress independently of vowel quality; and (4) allow for word endings to act as cues to stress across a variety of reading tasks.

As summarized throughout section 6.2, I would argue that my dissertation results best support the views advanced in Ševa and colleagues' (2009) connectionist model and in the sublexical route of CDP++ (Perry et al., 2010). Although the two models differ from each other in notable ways—for instance, in whether lexical representations are distributed (Ševa et al., 2009) or exist as a distinct route to word reading (Perry et al., 2010)—their claims are fairly similar with respect to the key issues that I explored in this dissertation. Specifically, both models suggest that readers learn of the associations between spelling and lexical stress through exposure to the regularities in text and that they subsequently use their understanding of those probabilistic regularities to place stress in written words. This learning applies across suffix and non-suffix word endings, captures the probabilistic nature of the links between stress and spelling, and (in the case of CDP++) can be separated from the influence of segmental phonology.

In drawing this theoretical conclusion, I am adding to the body of work that views readers' use of word endings as stress cues as being the result of connectionist processes, whereby statistical learning through exposure to text yields sensitivity to probabilistic links between stress and spelling. To date, this interpretation has been advanced to explain several empirical studies (e.g., Arciuli & von Koss Torkildsen, 2012; Colombo, Deguchi, & Boureux, 2014) and has been implemented as the basis for computational models (Arciuli et al., 2010; Perry et al., 2010; Ševa et al., 2009). Yet as I have outlined, the research presented in this dissertation supports and extends this existing literature by addressing previously untested theoretical claims and by exploring important gaps in our models of disyllabic word reading.

6.3 AN ACCOUNT OF READERS' SENSITIVITY TO, AND USE OF, ENDINGS AS STRESS CUES ACROSS DEVELOPMENT

Beyond their role in informing computational models, the findings presented in this dissertation can be used to further our understanding of multisyllabic word reading development. The connectionist and CDP++ models that I have reviewed thus far are aimed at skilled reading (Perry et al., 2010; Ševa et al., 2009; see Arciuli et al., 2010 for the sole attempt to model development), but both are built around a statistical learning mechanism that has developmental implications. In this section, I put forward a tentative account of the development in readers' use of word endings as cues to stress in multisyllabic written words, spanning early reading development through to skilled adult reading. The framework for this account is centered on connectionist principles, and I further explore and shape that framework by drawing on relevant empirical work, including the work in this dissertation.

Broadly, I argue that readers' use of word endings as stress cues changes based on two key factors. The first is the reader's underlying sensitivity to the statistical associations between stress and spelling in English. This sensitivity builds during reading development, emerging at different rates for different endings as a function of (a) the strength of an ending's association with stress, and (b) how widely that ending appears in text. By adulthood, sensitivity reaches a stable point that mirrors the regularities in the language. The second factor affecting readers' use of their sensitivity to word endings as stress cues is the reading task, and the demands it places on visual word recognition. More specifically, I suggest that the extent to which readers use endings as sublexical stress cues depends on the ease with which they can retrieve a word's stress pattern from the lexicon. Theoretical models (e.g., Perry et al., 2010) and empirical research (e.g., Protopapas, Panagaki, Andrikopoulou, Gutiérrez Palma, & Arvaniti, 2016) both suggest that stress patterns are stored as part of our lexical representations. Retrieving words from the lexicon therefore gives readers a deterministic source of information about stress (e.g. *COMet*), reducing the need to draw on probabilistic sublexical cues (e.g., *–et* → first syllable stress). I propose that characteristics of a reading task (e.g., isolated words vs. connected text) and its items (e.g., low vs. high frequency words; familiar words vs. pseudowords) may influence the ease with which stress can be retrieved from the lexicon, and, in turn, the extent to which word endings affect readers' performance. As a result, it is important that we consider task demands when exploring readers' developing sensitivity to word endings as stress cues.

When combined with prior empirical work, I submit that my dissertation's findings can speak to both of these factors, though only speculatively. Drawing conclusions about the development of readers' use of word endings as stress cues requires some caution, as the body of developmental work to date is small, cross-sectional, and methodologically varied. With that in mind, there is value in synthesizing the existing evidence, as doing so helps to highlight empirical gaps and explore theoretical predictions about stress and multisyllabic word reading development.

6.3.1 LEXICAL STRESS IN ORAL LANGUAGE DEVELOPMENT

Before children become readers, they build a rich foundation of oral language skills that will later form a basis for their reading development. For children in English-speaking environments, lexical stress is a key aspect of that foundation. Indeed, decades' worth of research shows that lexical stress has a widespread impact on oral language learning in early infancy and beyond. For instance, by 7 months of age, infants learning English use stress to segment words from continuous speech (e.g., Curtin, Mintz, & Christiansen, 2005). More specifically, they segment words that have first-syllable stress (e.g., *DOCtor*) and mis-segment words with second-syllable stress (e.g., *guiTAR*)—a tendency that reflects the fact that most words in English begin with a stressed syllable (Jusczyk, Houston, & Newsome, 1999). Infants also prefer to listen to words with the more common first-syllable stress pattern over those with the less common second-syllable stress pattern (Jusczyk, Cutler, & Redanz, 1993). By 12 months, infants can reliably learn and distinguish between words that differ only in their stress patterns

(Curtin, 2009), and they encode information about stress when learning new words even when stress is not needed to complete the task at hand (Curtin, 2011). Stress can also guide infants' processing of other features in oral language. As one example, a study of 20- to 24-month-olds found that the children could distinguish between words that differed by a single phoneme, but only did so successfully when the phoneme contrast occurred in the word's stressed syllable (Floccia, Nazzi, Austin, Arreckx, & Goslin, 2011). Taken together, these findings demonstrate that (a) infants who are learning English recognize stress as a salient feature of their language, and (b) they are sensitive to patterns in how stress is realized across words.

The role of lexical stress in spoken English persists among older infants and children, but it also evolves with their increasing exposure to the language. Much of that evolution involves changes in the broad preference for words with first-syllable stress. For example, unlike their 7-month-old counterparts, 10.5-month-old infants can segment words with second-syllable stress from continuous speech. This ability likely emerges through sensitivity to a variety of fine-grained word boundary cues that override the simple focus on first-syllable stress (Jusczyk et al., 1999). Similarly, while young infants are better able to learn words pronounced with first-syllable than second-syllable stress (e.g., Graf Estes & Bowen, 2013), the word learning advantage conferred to words with first-syllable stress disappears by 24 months (Floccia et al., 2011). We also see evidence that infants can overcome their broad bias toward words with first-syllable stress, capturing other regularities in the oral language. For example, 16-month-olds are only able to learn the labels for novel actions when those labels are pronounced with second-syllable stress (Curtin, Campbell, & Hufnagle, 2012). This suggests that even relatively young children recognize and use the link between stress and grammatical category in English—that is, most disyllabic verbs have second-syllable stress, whereas most disyllabic nouns have first-syllable stress (Kelly & Bock, 1988). These developmental changes are important, as they prepare children to contend with the wide variety of stress patterns that are possible in English words—not just words with the more common first-syllable stress pattern.

Notably, though, subtle differences remain in how preschool to school-aged children process and produce words with first- versus second-syllable stress. For instance, acoustic analyses of children's speech production show that their pronunciation of stress is adult-like by age 3 years, but only for words with first-syllable stress (Ballard et al., 2012). In contrast, their production of words with second-syllable stress follows a far more protracted developmental trajectory—among children as young as 3 years to those as old as 11 years, pronunciation of stress in words with second-syllable stress is not fully adult-like (Arciuli & Ballard, 2017; Ballard et al., 2012). As another example, oral language sensitivity to second-syllable stress is correlated with overall word reading ability among children aged 5–12 years, but the same is not true of their corresponding sensitivity to first-syllable stress (Arciuli, 2017). Finally, and most notably for the purposes of this dissertation, children in the early- and mid-elementary grades (ages 5 to 8 years) show a broad preference for first-syllable over second-syllable stress assignment when asked to pronounce disyllabic pseudowords (Arciuli et al., 2010).

Before considering how written word endings might act as cues to stress when reading, it is worth highlighting three salient points that emerge from the oral language literature reviewed in this section. First, stress is a core aspect of children's oral language learning, and is stored as a part of words' lexical representations from the age at which children's vocabularies begin to develop in earnest (approximately 12 months; Curtin, 2011). This is noteworthy for my purposes, because knowing a word's oral stress pattern is a crucial prerequisite to learning that there is an association between that stress pattern and the word's ending spelling. Second, the oral language literature shows that children have an early and persistent bias toward words with first-syllable stress—the most common stress pattern in English (e.g., Arciuli et al., 2010; Jusczyk et al., 1993). Children bring this oral language bias to their early encounters with text. As such, any written cues to stress will need to overcome children's existing preference for first-syllable stress. Finally, children's broad preference for first-syllable stress can be modulated by other oral language regularities (e.g., grammatical category; Curtin et al., 2012). I argue that written word endings can serve a similar modulating function once children begin to read text. Specifically, as children learn of the probabilistic associations between stress and ending spellings in English, they gradually come to use those word endings as fine-grained cues to stress placement in written words.

6.3.2 EARLY READING DEVELOPMENT

From the outset of reading development, children are equipped with the tools to start learning of the links between stress and spelling. Written word endings are associated with stress in children's earliest texts (Arciuli et al., 2010), and young children are sensitive to the stress patterns of words in oral language—a skill that is separable from their segmental phonological awareness (Beattie & Manis, 2014; Wade-Woolley & Heggie, 2016) and that builds through the early years of reading development (Calet, Gutiérrez-Palma, Simpson, González-Trujillo, & Defior, 2015; see also Goodman, Libenson, & Wade-Woolley, 2010; Holliman et al., 2017). Despite having both of these prerequisites in place, a connectionist account of the process by which readers learn of written stress cues would suggest that it takes time and ongoing text exposure for the full effects of those cues to emerge in children's reading behaviour.

Behavioural support for this gradually-emerging sensitivity comes from the small body of work on early readers' use of word endings as stress cues, which includes children through approximately the second grade. In a study of English-speaking children, for instance, Arciuli and colleagues (2010) showed that 5- and 6-year-old children made slight use of word endings to guide stress placement when reading pseudowords aloud—they gave words first-syllable stress roughly 12% more often when a pseudoword's ending cued first- rather than second-syllable stress. The same was true of 7- and 8-year-old children (a 19.9% difference). Importantly though, these young children's use of endings was overshadowed by their preference for first-syllable stress across all pseudowords. This pattern among English-speaking children has been corroborated with work on young readers of Italian; in pseudoword reading tasks, children show an early preference for the dominant stress pattern in the language that weakens with time (Colombo et al., 2014; Sulpizio & Colombo, 2013). Chapter 4 of this dissertation provides parallel findings with a different experimental paradigm—in it, children in the early-elementary grades showed no use of stress to guide their spelling choices. Instead, they strongly preferred the more common simple spellings for each ending (e.g., *–et*). Taken together, the behavioural work to date suggests a burgeoning but relatively weak sensitivity to the links between stress and spelling in English among early readers. I would argue that this early sensitivity is not yet strong enough to fully overcome children's existing preference for words with the dominant stress pattern in their oral language.

Missing from this early developmental picture is research on how word endings might affect readers' processing of familiar words. To date, the empirical work on stress cues in known words has focused on older readers (from Grade 4 and up; see below for a description of Sulpizio & Colombo, 2013, which provides an exception). While pseudoword tasks are a useful and straightforward way to assess whether readers recognize the links between stress and spelling in their language, it is also a situation where I would suggest that they are especially likely to draw on those cues. That is, pseudowords have no lexical influence for readers to draw on, almost requiring that children will make use of sublexical and/or distributional information to place stress. The same may not be true of children's real word reading, whether they encounter those words in isolation or in connected texts. As such, we do not yet have a full sense of the extent to which young readers use word endings during their everyday reading. Building on connectionist and statistical learning principles, we might not expect a particularly strong role of word endings among beginning readers, as their emerging sensitivity to the links between stress and spelling will not yet have fully developed.[22]

6.3.3 LATER READING DEVELOPMENT

By the mid-elementary grades, children begin to make more reliable use of word endings as cues to stress. This is consistent with both theory (e.g., Arciuli et al., 2010;

[22] Consistent with this speculation, Protopapas and Gerakaki (2009) showed that beginning Greek readers draw primarily on lexical sources of information when assigning stress (i.e., using the stress patterns of orthographically similar words to assign stress in Greek pseudowords). In fact, lexical sources of information had a stronger influence on stress placement than did diacritics, which provide explicit orthographic cues to stress.

Ševa et al., 2009) and the existing behavioural evidence (e.g., Arciuli et al., 2010; Colombo et al., 2014; Sulpizio & Colombo, 2013). From approximately age 9, we see that children focus less heavily on their language's dominant stress pattern than do their younger peers; instead, they increasingly use word endings to guide stress placement (Arciuli et al., 2010; Colombo et al., 2014; Sulpizio & Colombo, 2013). In Arciuli and colleagues' (2010) study of English-speaking children, 9- to 12-year olds made significantly greater use of endings to guide pseudoword stress placement than their younger peers—they assigned first-syllable stress to pseudowords whose endings cued first syllable stress 36% more often (9- to 10-year-olds) to 42% more often (11- to 12-year-olds) than they did words whose endings cued second-syllable stress. And notably, these older children had no strong general bias toward English's dominant first-syllable stress pattern; they preferentially gave pseudowords second-syllable stress when the word's spelling called for it (see Colombo et al., 2014; Sulpizio & Colombo, 2013 for similar findings in Italian). Broadly, this cross-sectional trend aligns with the connectionist claim that older readers will be more sensitive to the links between stress and spelling than younger readers.

However, adding my dissertation's findings to this prior work yields a mixed picture of the point at which developing readers first show a reliable sensitivity to word endings' associations with stress. In Chapter 4, for instance, children in the mid-elementary grades made no use of stress to guide their pseudoword spelling choices. The effect of stress on spelling did not emerge until the late-elementary grades, when children were approximately 10 to 12 years old. This is somewhat later than the age at which children in prior studies have used ending spellings to assign stress in pseudowords

(Arciuli et al., 2010; Sulpizio & Colombo, 2013). I would argue that, in part, this discrepancy can be explained by a core tenet of statistical learning theory: that readers' sensitivity to the various regularities in text emerges gradually, through accumulated exposure to text input (e.g., Perruchet & Pacton, 2006). Crucially, this exposure will happen more quickly for some endings than for others as a function of the strength of endings' associations with stress and how commonly a given ending occurs in text. In other words, during reading development, children should be more sensitive to some endings' role as stress cues than others'. In prior work, Arciuli and colleagues (2010) used ending spellings that are relatively common in children's texts (e.g., *–et*, *–act*), thereby evaluating children's sensitivity to ending cues to which they would have had ample exposure. In contrast, the spelling choice task used in Chapter 4 included endings that are comparatively rare in children's texts (e.g., *–ette*; see Appendix H). In doing so, I suggest that my study may have tested the upper limits of mid-elementary aged children's sensitivity to word endings as stress cues, yielding the null effect of stress on spelling choices until the late-elementary grades.

My interpretation of this discrepancy is certainly speculative—testing this conclusion will require more direct evidence than a comparison of methodologically different cross-sectional studies. However, the design used in Chapter 3 would be a useful way to explore my claim that young readers' sensitivity to word endings as stress cues builds gradually, emerging sooner for the endings that children encounter frequently in texts. That is, in a naming and/or lexical decision task, children would be presented with a range of words whose endings vary in (a) the strength of their association with stress, and (b) how often they occur in children's texts. If my speculation is correct, I would expect the predictions I made in Chapter 3 to be borne out in developing readers—the word endings that they encounter most often should have the strongest effect on their word reading and recognition. By examining this possibility in children, whose sensitivity to word endings as stress cues is still emerging, we will be better able to address key predictions of the statistical learning framework.

Furthermore, as with younger children, I submit that it is important to consider how task demands might affect the use of word endings as stress cues. The study that I presented in Chapter 5 speaks to this point. In it, word endings' consistency with stress had no effect on children's processing of words in connected text (when using endings that are relatively common in children's texts; Appendix H). This was true across the mid- and late-elementary grades, which, notably, are grades during which children use the links between stress and spelling to complete pseudoword tasks (Chapter 4; Arciuli et al., 2010; Sulpizio & Colombo, 2013). This indicates an underlying sensitivity to the links between stress and spelling. Yet the discrepancies between my Chapter 5 findings and that of prior work further suggests that children may apply their sensitivity selectively across tasks. Earlier, I proposed that ease of lexical access might influence whether—and how much—readers draw on sublexical cues to stress like word endings. This interpretation would suggest greater use of endings when processing pseudowords (which lack lexical representations) than when processing known words, and when processing words in isolation rather than in connected text.[23] Certainly, more research is

[23] In support of this possibility, we can refer to the small body of research on Italian-speaking children's isolated word reading. Studies with children in Grade 4 (Burani et al., 2014) and in Grade 6 (Paizi, Zoccolotti, & Burani, 2011; typical readers) showed that children made fewer stress placement errors in words whose endings accurately cued their stress patterns. Both studies focused specifically on reading of low frequency words

needed on this question. My dissertation's studies contribute evidence from two novel tasks, adding to the limited work on English-speaking children's sensitivity to word endings as cues to stress (cf. Arciuli et al., 2010), but they do not provide the direct task comparisons that would act as more stringent tests of my task demand predictions. Even so, the discrepant findings across tasks serve as an important reminder that it is not enough to establish children's sensitivity to word endings as stress cues; we must also consider the extent to which they use those cues in typical reading tasks.

6.3.4 SKILLED ADULT READING

The earlier sections of this discussion have largely outlined the theoretical contributions of my findings in adult readers. In a developmental context, I see these findings as an extension of the trend I outlined for developing readers—just as endings have a stronger effect on mid- to late-elementary aged students than younger ones (e.g., Arciuli et al., 2010; Sulpizio & Colombo, 2013), endings have a stronger effect on adults than on children (e.g., Chapter 4; Chapter 5; Burani, Paizi, & Sulpizio, 2014; Colombo et al., 2014). In this respect, skilled adult reading reflects an endpoint in the connectionist process by which readers learn of the associations between stress and spelling in English. When combined with prior research, my dissertation's findings show that adults use word endings as cues to stress in written words across a wide variety of tasks—to place stress in pseudowords (e.g., Arciuli & Cupples, 2006) and guide pseudoword spelling decisions (Chapter 4), to recognize words in isolation (Chapter 2; Chapter 3; Kelly et al., 1998;

---

in an effort to limit lexical influences. Other findings give preliminary support for this focus on low frequency words—Sulpizio and Colombo (2013) report an effect of endings' consistency with stress on children's reading of low frequency, but not high-frequency, words (though this interaction should be interpreted cautiously, as ending consistency was confounded with stress dominance).

Mundy & Carroll, 2013), and even when reading words in connected text (Chapter 5). This speaks to a robust role for word endings as stress cues in English-speaking adults. And notably, this role for word endings as stress cues is not limited to tasks that involve reading words aloud (as is the focus of the existing models of disyllabic word reading; Arciuli et al., 2010; Perry et al., 2010; Rastle & Coltheart, 2000; Ševa et al., 2009). Instead, the data suggest that adults draw on cues to stress in written words even when the reading task requires no phonological output.

I would argue that my dissertation's findings (specifically Chapter 3's) show a stability in adults' underlying sensitivity to the links between stress and spelling in English. That is, through years of exposure to text, they have encountered enough examples of all word endings—including relatively rare ones—to have an accurate sense of the probabilistic regularities that exist in the language. From a connectionist standpoint, this would mirror the point at which a model's weightings stabilize during training, where further exposure to words yields diminishing returns. This interpretation would explain why adults in Chapter 3 were equally sensitive to word endings that appear in relatively few words and those that appear in many. With this in mind, I see adulthood as an ideal point at which to systematically explore the role of task demands on readers' use of endings as stress cues. The patterns across studies to date are suggestive—for example, in Chapter 5 I showed that endings' effect on words read in text was smaller than that on words read in isolation (e.g., Mundy & Carroll, 2013). That said, research that directly explores task differences will provide more compelling tests of the circumstances under which skilled adult readers use word endings as cues to stress in multisyllabic written words.

6.4 LIMITATIONS AND FUTURE DIRECTIONS

6.4.1 OPERATIONALIZING ENDINGS AS STRESS CUES

Conducting research on the role of word endings as stress cues requires that we operationally define written word endings. Throughout the work in this dissertation, I adopted the definition proposed by Arciuli and Cupples (2006), which identifies word endings as written units that map onto the rime of a word's second syllable. As I noted in the General Introduction, there are certainly benefits to the use of this definition: it is precise, can be applied systematically to words in the English lexicon (Arciuli & Cupples, 2006), and it maps onto a psycholinguistic unit that is theoretically relevant to word reading (e.g., Ziegler & Goswami, 2005). Taken together, the strengths of Arciuli and Cupples' (2006) definition allow us to establish the scope and prevalence of word endings' association with stress in a way that is linguistically motivated and that avoids certain practical pitfalls. For these reasons, I argue that it is an effective working definition with which to capture a statistical association between stress and spelling in English.

Despite those strengths, it is not yet clear if rime units are the only—or indeed the best—way to characterize written word endings as stress cues in English. In principle, the associations between spelling and stress that are captured through rime units might be tapping into a different, overlapping written cue. For example, readers might draw on the spelling of a word's final phoneme (e.g., *–tte* in *rouLETTE*) or its final syllable (e.g., *–lette*) as the salient stress cue, rather than its final rime (e.g., *–ette*). Resolving this issue will require corpus and behavioural research in which a range of linguistically-motivated units are compared on the extent to which they impact readers' processing of lexical stress in written words. To date, the only such comparison has come from work that contrasts the roles of word endings and beginnings as stress cues (e.g., Arciuli et al., 2010; Monaghan, Arciuli, & Ševa, 2016; see also Kelly, 2004). In these studies, word beginnings are defined as the body of a word's first syllable (i.e., the letters up to and including the first phonemic vowel; Arciuli & Cupples, 2007), and word endings remain defined as the rime of the final syllable. Corpus analyses show that both are associated with stress in English, but that endings are the more reliable cue in both adults' and children's texts (Arciuli et al., 2010; Monaghan et al., 2016). Behaviourally, children show increasing sensitivity with age to word endings over beginnings as stress cues (Arciuli et al., 2010). These findings support the claim that word endings are useful as written stress cues. Still, they leave questions about the best definition of word endings unanswered.

A recently-published chapter by Monaghan and colleagues (2016) acknowledges this issue and begins to explore it with a systematic corpus analysis. In it, they compared a range of word endings and word beginnings on how accurately each cued stress in several languages, English among them. Here, I focus on their analysis of various ending units in disyllabic English words, as those are the results that most closely relate to my dissertation. Three of the ending units that Monaghan and colleagues explored were linguistically motivated (the word's final rime, final consonant, and final vowel, respectively), while the others were based on a specified number of letters (the last one to five letters of a word, respectively). Overall, their findings affirm that rime-based ending units are useful as probabilistic stress cues. Rimes classified stress more accurately than either of the other linguistically motivated units, and encouragingly, rimes also outperformed the blind use of the last one or two letters in a word.[24] These analyses are certainly a step in the right direction, but they do not fully resolve the issue of how to best operationalize word endings as stress cues (as the authors themselves acknowledge). First, Monaghan and colleagues focused solely on endings' ability to classify stress accurately; they did not aim to identify the optimal unit for determining stress placement. Secondly, they compared only a small number of linguistically-motivated ending units; others might turn out to be more useful than rimes. Most importantly, their work was corpus-based, and thus cannot speak to the ending cues that most affect readers' processing of multisyllabic written words. As such, behavioural work that builds on Monaghan and colleagues' analyses will be informative.

Although the question of how to best define written word endings is important to fully understanding the role of endings as stress cues, I would suggest that it is more of a methodological issue than a theoretical one. Certainly, interpreting my dissertation findings under a rule-based model for stress assignment would require that a precise definition of word endings be specified *a priori* (e.g., Rastle & Coltheart, 2000). However, connectionist models are well-suited to accommodating situations where there is ambiguity in the cues most relevant to word reading. As Ševa and colleagues (2009) put it, connectionist models allow for stress assignment to emerge "through a

[24] Unsurprisingly, Monaghan and colleagues (2016) found that stress classification accuracy increased as did the number of letters in an ending unit. Indeed, units based on the last four and five letters of words classified stress more accurately than did rimes. When interpreting these results, it is worth noting that capturing more information in a word (in the form of more letters) necessarily yields more accurate classifications (Monaghan et al., 2016). The majority of English rime-based endings are three letters long, and many more are two letters; relatively few are four or five letters long (Arciuli & Cupples, 2006; Baayen et al., 1995). This helps to explain the high stress classification accuracy of units that comprised a word's last four and five letters (Monaghan et al., 2016).

combination of different cues present in the orthographic input … from single letters, bigrams, trigrams, and so on … [that] do not have to be specified or listed in advance." (p. 246). In short, it is possible that the rime units used in my research are not the optimal way to quantify word endings as stress cues in English. However, even if that turns out to be true, a model developed under the connectionist framework should still be able to discover the underlying association between stress and spelling that my ending definition is tapping into.

6.4.2 INDIVIDUAL DIFFERENCES IN SENSITIVITY TO STRESS CUES

Thus far, I have made a case that exposure to the regularities in text is what drives readers' sensitivity to written word endings as stress cues in English. In doing so, I have focused on word-level indices of exposure, using variables that capture differences in how reliably word endings act as cues to stress and how often words with a given ending cue appear. These item-level differences were the basis for each of the experiments in this dissertation, and they are the focus of the theoretical models that inspired their research questions (Rastle & Coltheart, 2000; Perry et al., 2010; Ševa et al., 2009).

However, differences in exposure to text regularities are not just a function of word-level characteristics; individual readers also differ in how much text they encounter (e.g., Acheson, Wells, & MacDonald, 2008; Moore & Gordon, 2015; Stanovich & West, 1989). These individual differences are conceptually important to models of word reading, particularly when viewing readers' sensitivity to text regularities from a connectionist standpoint (though neither existing connectionist model attempted to simulate reader differences in text exposure; Perry et al., 2010; Ševa et al., 2009). That is, individuals who read widely and often will have more opportunities to learn of all regularities in text—including the links between stress and spelling—than their peers with less print exposure. We might therefore expect readers with high levels of print exposure to be particularly sensitive to those regularities.

A study by Falkauskas and Kuperman (2015) provides a nice demonstration of this point. Their study explored readers' processing of compound phrases, and showed that adults have a processing bias that favours the more common spelling variant of each compound over the less common spelling (e.g., *lunchroom* over *lunch room*; *boot camp* over *bootcamp*). Importantly, this bias was modulated by two key factors, which interacted to affect readers' processing of the words: the overall frequency of each compound (an item-level difference in exposure to text regularities) and the participant's level of print exposure (a reader-level difference in exposure to all text regularities). In other words, less experienced readers were less sensitive to the statistical regularities in text than were more experienced readers. This finding highlights the value in capturing individual differences in work that seeks to understand how the characteristics of written language affect word reading and recognition.

Individuals also differ in their oral language vocabularies, and more specifically in their knowledge of words' pronunciations. Like print exposure, these differences may impact readers' sensitivity to written word endings as stress cues—after all, without knowing words' oral stress patterns, readers cannot link stress with its written correlates. While this is certainly a conceptual issue for my dissertation research, the words used in my tasks (both as items and as the basis for quantifying endings' links with stress) are ones that participants are likely to be familiar with in both oral and written language. This fact is largely thanks to my focus on disyllabic words (rather than longer words that are more likely to be rare in spoken language). A review of the words that I used with adults suggests that excluding particularly rare words (e.g., *rennet*; *alee*; *luddite*) does not systematically affect my calculation of ending characteristics, nor does it impact my item selection (CELEX database; Baayen et al., 1995). Similarly, the words that I used in my experiments with children are generally more common in young children's oral language input (from age 1;0 to 7;9, based on the ChildFreq database; Bååth, 2010) than in their written language input (during the elementary grades; Zeno et al., 1995). Certainly, it will be useful for future research to verify the assumption that participants know target words in both oral and written language. In addition to quantifying words' oral language frequencies, a fruitful way to explore this issue might be to include measures of readers' spoken vocabulary.

As such, future work that builds on the findings in this dissertation might consider how differences in readers' print exposure and/or oral language vocabulary affect their sensitivity to—and use of—word endings as stress cues. Broadly, I would expect differences in both skills to affect the timing of the developmental sequence that I proposed in section 6.3. That is, I would expect all readers to (a) gradually acquire sensitivity to different word endings as stress cues through the early- to mid-elementary grades, (b) initially be more sensitive to endings that appear often than those that are uncommon, and (c) eventually accumulate enough exposure to text that their sensitivity to all stress cues is stable, with further exposure yielding diminishing behavioural returns. However, I would expect those milestones to emerge later among readers with relatively low levels of print exposure than among readers who have more print exposure, and among those with smaller than larger oral language vocabularies. Empirical support for this prediction would help to validate my claim that exposure drives readers' sensitivity to endings as stress cues, making it a valuable avenue for future research.

Although my dissertation work may have benefitted from considering differences in readers' print exposure, lacking these measures does not fundamentally affect my interpretation of the findings. Support for this perspective comes from the crossed-random effects of the analyses presented in Chapter 3, which explored the effect of words' ending characteristics on readers' lexical decision latencies. These were my only analyses to directly incorporate potential reader-level differences. Specifically, the models included random slopes for all item-level variables (e.g., word frequency, proportion of stress friends, etc.). This allows for the effect of each variable to differ for each participant—one reader might have quite a steep slope for the proportion of stress friends variable, suggesting that it has a strong effect on their processing of written words; in contrast, another reader's slope on that variable might be shallow, suggesting a weaker effect. The random slopes for the three variables characterizing words' endings are of particular interest for my purposes (proportion of stress friends; number of stress friends; number of stress enemies). Descriptively, there were indeed slight differences in the slopes of these variables across participants. But importantly, those random slope differences did not contribute at all to the models (accounting for 0% variance in the reported models). In other words, the effects of word endings as stress cues was comparable across all participants of the study reported in Chapter 3. These negligible random slope differences suggest that my item-level focus was not particularly problematic for those models. However, it is worth noting that the study was not designed to explore differences between readers. As such, its findings should not be used to discount the possibility that individual difference variables might affect readers' sensitivity to word endings as stress cues.

In addition to print exposure, there are a variety of individual difference measures that might affect readers' ability to learn of the links between stress and spelling in English. Two stand out as particularly relevant: domain-general statistical learning ability and prosodic awareness. Interestingly, both skills have been linked with word reading ability. For instance, research by Arciuli and Simpson (2012) shows that both adults and school-aged children are able to implicitly learn of regularities embedded in a string of presented images at above-chance levels. However, individuals' accuracy in identifying regularities varies widely, and these differences predict a significant, albeit modest, amount of variance in word reading accuracy—those with stronger statistical learning abilities tended to be better readers. Similarly, prosodic awareness, which includes the ability to recognize the stress patterns in oral language, predicts word reading among children and adults. For instance, children who are better able to match a word or phrase with a presented stress pattern (which has been stripped of phonemic information) tend to outperform their peers on standardized measures of general word reading (e.g., Arciuli, 2017; Clin, Wade-Woolley, & Heggie, 2009; Goswami, Gerson, & Astruc, 2010) and on more targeted measures of multisyllabic word reading (Holliman et al., 2017). Adults' ability to identify and manipulate the stress patterns of aurally presented words likewise predicts their word reading success (Chan & Wade-Woolley, 2016).

Beyond these links with word reading, I would suggest that both of these skills are directly relevant to the ability to learn of the role of word endings as cues to stress in written words. Readers with stronger statistical learning abilities are equipped to apply that domain-general skill to the learning of specific regularities, including the associations between word endings and lexical stress. Similarly, readers can only learn of the associations between spelling and stress if they are able to effectively recognize stress patterns in oral language. As such, we might expect a reader with relatively strong prosodic awareness to be particularly well-equipped to linking stress with its orthographic cues; for a reader with more limited awareness of aural stress patterns, learning of these associations is likely to be a challenge. Individual differences in statistical learning ability and prosodic awareness therefore offer promising avenues for future research that seeks to explore the abilities behind readers' sensitivity to word endings as cues to stress in written words.

6.5 CONCLUSIONS

Lexical stress is central to the phonology and lexical identity of multisyllabic words in English, which comprise the majority of words in the lexicon (Baayen, Piepenbrock, & van Rijn, 1995). Despite its importance, placing stress in English words is not trivial matter—stress patterns vary across words and have no direct orthographic markings. With these challenges in mind, understanding how readers process lexical stress in written words is paramount to understanding multisyllabic word reading more generally. To that end, the research that I have presented in this dissertation adds to the growing evidence that skilled and developing readers of English use written word endings as probabilistic cues to stress (e.g., Arciuli & Cupples, 2006; Arciuli et al., 2010; Kelly et al., 1998; Mundy & Carroll, 2013; Verrekia, 1996). My findings build on this prior work by offering novel insight into the nature of readers' sensitivity to word endings as stress cues and by differentiating between competing theoretical accounts of disyllabic word reading (Perry et al., 2010; Rastle & Coltheart, 2000;Ševa et al., 2009).

When taken together, the experiments in this dissertation speak to a robust role for written word endings as cues to stress—one that applies across a range of ending types, stress neighbourhoods, and reading tasks. The findings that I have presented show that word endings' associations with stress affect readers similarly regardless of whether or not those endings can serve as English suffixes (Chapter 2) and regardless of whether the association is demonstrated by a large or small number of stress friends (Chapter 3). They offer support for the claim that readers' sensitivity to written stress cues emerges from their relative exposure to all members of a stress neighbourhood—both stress friends and stress enemies (Chapter 3). This dissertation's findings also validate the role of stress itself in the processing of written words by showing that experienced readers' sensitivity to the links between word endings and lexical stress persists when vowel quality is controlled (Chapter 4) and suggest that word endings affect skilled readers' processing of words in connected text (Chapter 5). Finally, my findings speak to developmental trends in young readers' sensitivity to word endings as stress cues. Despite evidence that even young children use word endings to guide stress placement (Arciuli et al., 2010), it is not until the late-elementary grades that readers reliably use stress to guide ending spelling decisions (Chapter 4), and not until adulthood that word endings affect their reading of words in connected text (Chapter 5). I argue that these conclusions align with a connectionist view of stress placement in multisyllabic word reading (Perry et al., 2010; Ševa et al., 2009). According to this theoretical perspective, cumulative exposure to text yields a gradually-emerging sensitivity to probabilistic links between stress and spelling.

This interpretation coherently explains many of the findings in this dissertation, and notably, it provides a basis from which further research can explore the factors that affect readers' developing sensitivity to word endings as cues to stress. Overall, this dissertation helps to highlight the important role of lexical stress in processing written words, and as such, it represents a valuable, incremental step toward a better understanding of multisyllabic word reading in English.

6.6 APPENDIX H: ELEMENTARY GRADE FREQUENCIES OF WORD ENDINGS USED AS STRESS CUES IN RESEARCH

In this appendix, I compare elementary grade-level frequencies of the word endings used in studies of English-speaking children's sensitivity to endings as stress cues. Specifically, I compare frequencies at Grades 1 through 6 for: (1) the simple ending spellings used in Chapter 4, Experiment 2 [pseudoword spelling choice paradigm]; (2) the extended ending spellings used in Chapter 4, Experiment 2 [pseudoword spelling choice paradigm]; (3) the endings used in Chapter 5, Experiment 2 [Missing Letter Effect (MLE) paradigm]; and (4) the endings used in Arciuli, Monaghan, and Ševa's (2010) study of pseudoword stress assignment. Each set of endings is listed in the table below.

Chapter 4 (simple spellings)	Chapter 4 (extended spellings)	Chapter 5 (MLE study)	Arciuli et al. (2010)
–eak	*–ique*	*–end*	*–act*
–et	*–ette*	*–ent*	*–ade*
–o	*–eau*	*–ice*	*–an*
–um	*–umb*	*–ist*	*–et*
–us	*–uss*	*–ose*	*–ey*
–y	*–ee*	*–uce*	*–ince*
		–ume	*–ol*
		–ure	*–oon*

I quantify ending frequencies in two ways: by tallying the total number of words in which endings appear at each grade level (see Figure 6.1), and by computing the summed frequencies of the words with each ending at each grade level (see Figure 6.2). All frequencies are based on grade-level corpora from *The Educator's Word Frequency Guide* (Zeno et al., 1995).

Two comparisons inform the interpretations that I put forward in section 6.3.2 of this dissertation. First, the endings in Arciuli and colleagues' (2010) study appear significantly more often in children's texts than do Chapter 4's extended ending spellings. I use this difference to explain why children use endings to guide stress placement from the early- and mid-elementary grades (Arciuli et al., 2010), but do not use stress to guide ending spelling decisions until the late-elementary grades (Chapter 4). I argue that children's sensitivity to word endings as stress cues emerges gradually through exposure to the regularities in text. As such, young readers may not recognize that rare endings (e.g., *–ette*, as in Chapter 4) are associated with stress despite being sensitive to more common endings' associations with stress (e.g., *–act*, as in Arciuli et al., 2010).

The second relevant comparison is between the endings in Arciuli and colleagues' study and those used in Chapter 5. At all grade levels, these two sets of endings appear in a similar number of words that have similar summed frequencies. As a result, differences in ending frequency cannot explain the fact that word endings affect young children's pseudoword stress assignment (Arciuli et al., 2010) but do not affect children's letter detection rates when reading words in connected text (Chapter 5). This led me to consider alternate explanations for the discrepancy across studies. In section 6.3.2, I suggest that task demands (e.g., assigning stress to isolated pseudowords vs. reading familiar words in connected text) may affect the extent to which children draw on endings as sublexical cues to stress in written words.

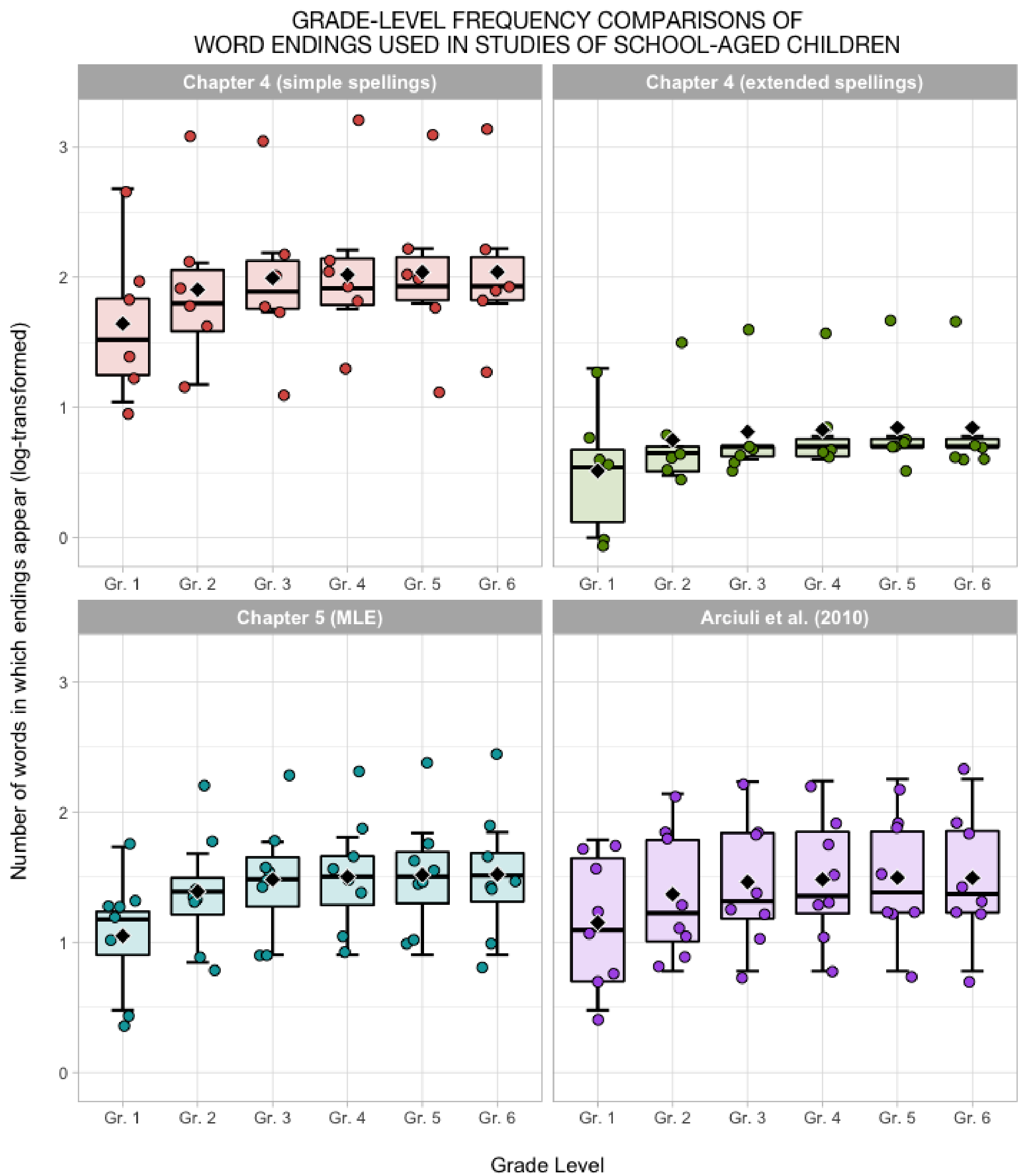


Figure 6.1. Boxplots and scatterplots showing grade-level frequencies of the word endings used in studies of English-speaking children's sensitivity to endings as stress cues. Frequency is quantified as the number of words in which each ending appears. At every grade level, the extended spellings used in Chapter 4 appear in fewer words than those used in Arciuli and colleagues' (2010) pseudoword naming study (*p*s ≤ .043). However, Chapter 5's MLE endings and the endings used in Arciuli and colleagues' (2010) study appear in a similar number of words (*p*s ≥ .789). Scatterplot data points represent the frequency of a given ending. Diamonds represent the mean frequencies at each grade level.

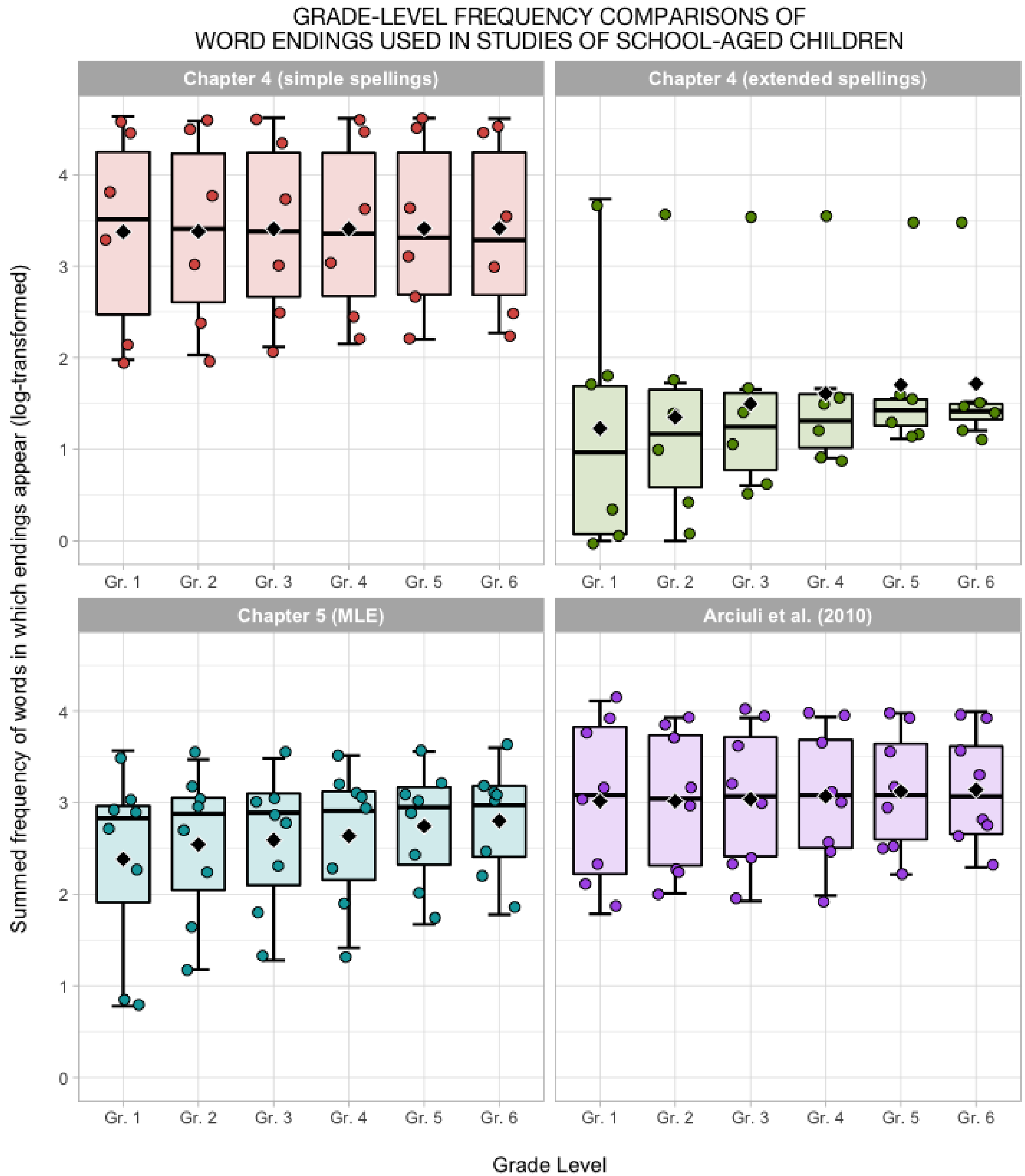


Figure 6.2. Boxplots and scatterplots showing grade-level frequencies of the word endings used in studies of English-speaking children's sensitivity to endings as stress cues. Frequency is quantified as the summed frequency of words in which each ending appears. At every grade level, the extended spellings used in Chapter 4 appear less frequently than those used in Arciuli and colleagues' (2010) pseudoword naming study (*ps* $\leq$ .013). However, Chapter 5's MLE endings and the endings used in Arciuli and colleagues' (2010) study appear with similar frequency (*ps* $\geq$ .234). Scatterplot data points represent the frequency of a given ending. Diamonds represent the mean frequencies at each grade level.

REFERENCES

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods, 40*(1), 278–289. http://dx.doi.org/10.3758/BRM.40.1.278

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

Adelman, J. S. (2012). Methodological issues with words. In J. S. Adelman (Ed.), *Visual word recognition, Volume I: Models and methods, orthography and phonology* (pp. 116–138). New York, NY: Psychology Press.

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review, 105*(4), 678–723. http://dx.doi.org/10.1037/0033-295X.105.4.678-723

Archer, A. L., Gleason, M. M., & Vachon, V. L. (2003). Decoding and fluency: Foundation skills for struggling older readers. *Learning Disability Quarterly, 26*(2), 89–101. http://dx.doi.org/10.2307/1593592

Arciuli, J. (2017). The relationship between children's sensitivity to dominant and non-dominant patterns of lexical stress and reading accuracy. *Journal of Experimental Child Psychology, 157*, 1–13. http://dx.doi.org/10.1016/j.jecp.2016.11.016

Arciuli, J., & Ballard, K. J. (2017). Still not adult-like: Lexical stress contrastivity in word productions of eight- to eleven-year-olds. *Journal of Child Language, 44*(5), 1274–1288. http://dx.doi.org/10.1017/S0305000916000489

Arciuli, J., & Cupples, L. (2006). The processing of lexical stress during visual word recognition: Typicality effects and orthographic correlates. *The Quarterly Journal of Experimental Psychology, 59*(5), 920–948. http://dx.doi.org/10.1080/02724980443000782

Arciuli, J., & Cupples, L. (2007). Would you rather 'embert a cudsert' or 'cudsert an embert'? How spelling patterns at the beginning of English disyllables can cue grammatical category. In A. Schalley, & D. Khlentzos (Eds.), *Mental states: Language and cognitive structure* (pp. 213–237). Amsterdam: John Benjamins.

Arciuli, J., Monaghan, P., & Ševa, N. (2010). Learning to assign lexical stress during reading aloud: Corpus, behavioral, and computational investigations. *Journal of Memory and Language, 63*(2), 180–196. http://dx.doi.org/10.1016/j.jml.2010.03.005

Arciuli, J., & Paul, R. (2012). Sensitivity to probabilistic orthographic cues to lexical stress in adolescent speakers with autism spectrum disorder and typical peers. *The Quarterly Journal of Experimental Psychology, 65*(7), 1288–1295. http://dx.doi.org/10.1080/17470218.2012.655700

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science, 36*(2), 286–304. http://dx.doi.org/10.1111/j.1551-6709.2011.01200.x

Arciuli, J., & von Koss Torkildsen, J. (2012). Advancing our understanding of the link between statistical learning and language acquisition: The need for longitudinal data. *Frontiers in Psychology, 3*(324), 1–9. http://dx.doi.org/10.3389/fpsyg.2012.00324

Ashby, J., & Clifton, C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition, 96*(3), B89–B100. http://dx.doi.org/10.1016/j.cognition.2004.12.006

Bååth, R. (2010). ChildFreq: An online tool to explore word frequencies in child language. *LUCS Minor, 16,* 1–6.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* New York, NY: Cambridge University Press.

Baayen, R. H., Bates, D. M., Kliegl, R., & Vasishth, S. (2015). *RePsychLing: Data sets from psychology and linguistics experiments* (Version 0.0.4) [R Package]. Retrieved from https://github.com/dmbates/RePsychLing

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412. http://dx.doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*(2), 12–28. http://dx.doi.org/10.21500/20112084.807

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1995). *The CELEX lexical database* (Version 2) [Computer Software]. Philadelphia, PA: Linguistic Data Consortium.

Ballard, K. J., Djaja, D., Arciuli, J., James, D. G. H., & van Doorn, J. (2012). Developmental trajectory for production of prosody: Lexical stress contrastivity in children ages 3 to 7 years and in adults. *Journal of Speech, Language, and Hearing Research, 55*(6), 1822–1835. http://dx.doi.org/10.1044/1092-4388(2012/11-0257)

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods, 39*(3), 445–459. http://dx.doi.org/10.3758/BF03193014

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. http://dx.doi.org/10.1016/j.jml.2012.11.001

Barry, C., & Seymour, P. H. K. (1988). Lexical priming and sound-to-spelling contingency effects in nonword spelling. *The Quarterly Journal of Experimental Psychology, 40*(1), 5–40. http://dx.doi.org/10.1080/14640748808402280

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. *arXiv Preprint.* https://arxiv.org/abs/1506.04967

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2017). *lme4* (Version 1.1-12) [R Package]. Retrieved from https://CRAN.R-project.org/package=lme4

Beattie, R. L., & Manis, F. R. (2014). The relationship between prosodic perception, phonological awareness and vocabulary in emergent literacy. *Journal of Research in Reading, 37*(2), 119–137. http://dx.doi.org/10.1111/j.1467-9817.2011.01507.x

Boersma, P., & Weenink, D. (2017). *Praat: Doing phonetics by computer* (Version 6.0.25) [Computer Software]. Retrieved from http://www.praat.org/

Bond, Z. S., & Small, L. H. (1983). Voicing, vowel, and stress mispronunciations in continuous speech. *Attention, Perception, & Psychophysics, 34*(5), 470–474. http://dx.doi.org/10.3758/BF03203063

Brand, A., & Bradley, M. T. (2016). The precision of effect size estimation from published psychological research: Surveying confidence intervals. *Psychological Reports, 118*(1), 154–170. http://dx.doi.org/10.1177/0033294115625265

Breen, M., & Clifton, C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language, 64*(2), 153–170. http://dx.doi.org/10.1016/j.jml.2010.11.001

Bryant, D. P., Ugel, N., Thompson, S., & Hamff, A. (1999). Instructional strategies for content-area reading instruction. *Intervention in School and Clinic, 34*(5), 293–302. http://dx.doi.org/10.1177/105345129903400506

Burani, C., & Arduino, L. S. (2004). Stress regularity or consistency? Reading aloud Italian polysyllables with different stress patterns. *Brain and Language, 90*(1), 318–325. http://dx.doi.org/10.1016/S0093-934X(03)00444-9

Burani, C., Paizi, D., & Sulpizio, S. (2014). Stress assignment in reading Italian: Friendship outweighs dominance. *Memory & Cognition, 42*(4), 662–675. http://dx.doi.org/10.3758/s13421-013-0379-5

Calet, N., Gutiérrez-Palma, N., Simpson, I. C., González-Trujillo, M. C., & Defior, S. (2015). Suprasegmental phonology development and reading acquisition: A longitudinal study. *Scientific Studies of Reading, 19*(1), 51–71. http://dx.doi.org/10.1080/10888438.2014.976342

Carlisle, J. F., & Stone, C. A. (2005). The role of morphemes in word reading. *Reading Research Quarterly, 40*(4), 428–449. http://dx.doi.org/10.1598/RRQ.40.4.3

Carlisle, J. F., Stone, C. A. and Katz, L. A. (2001) The effects of phonological transparency on reading derived words, *Annals of Dyslexia, 51*(1), 249–274. http://dx.doi.org/10.1007/s11881-001-0013-2

Casalis, S., Quémart, P. & Duncan, L. G. (2015). How language affects children's use of derivational morphology in visual word and pseudoword processing: Evidence from a cross-language study. *Frontiers in Psychology, 6*(452), 1–10. http://dx.doi.org/10.3389/fpsyg.2015.00452

Chall, J. S. (1983). *Stages of reading development*. New York, NY: McGraw-Hill.

Chan, J. S., & Wade-Woolley, L. (2016). Explaining phonology and reading in adult learners: Introducing prosodic awareness and executive functions to reading ability. *Journal of Research in Reading.* Advance online publication. http://dx.doi.org/10.1111/1467-9817.12083

Clin, E., Wade-Woolley, L., & Heggie, L. (2009). Prosodic sensitivity and morphological awareness in children's reading. *Journal of Experimental Child Psychology, 104*(2), 197–213. http://dx.doi.org/10.1016/j.jecp.2009.05.005

Colombo, L. (1992). Lexical stress effect and its interaction with frequency in word pronunciation. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 987–1003. http://dx.doi.org/10.1037/0096-1523.18.4.987

Colombo, L., Deguchi, C., & Boureux, M. (2014). Stress priming and statistical learning in Italian nonword reading: Evidence from children. *Reading and Writing: An Interdisciplinary Journal, 27*(5), 923–943. http://dx.doi.org/10.1007/s11145-013-9476-x

Colombo, L., & Sulpizio, S. (2015). When orthography is not enough: The effect of lexical stress in lexical decision. *Memory & Cognition, 43*(5), 811–824. http://dx.doi.org/10.3758/s13421-015-0506-6

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review, 100*(4), 589–608. http://dx.doi.org/10.1037/0033-295X.100.4.589

Coltheart, M., Davelaar, E., Jonasson, T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Lawrence Erlbaum Associates.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256. http://dx.doi.org/10.1037/0033-295X.108.1.204

Cunningham, P. M. (1998). The multisyllabic word dilemma: Helping students build meaning, spell, and read "big" words. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 14*(2), 189–218. http://dx.doi.org/10.1080/1057356980140204

Curtin, S. (2009). Twelve-month-olds learn novel word–object pairings differing only in stress pattern. *Journal of Child Language, 36*(5), 1157–1165. http://dx.doi.org/10.1017/S0305000909009428

Curtin, S. (2011). Do newly formed word representations encode non-criterial information? *Journal of Child Language*, *38*(4), 904–917. http://dx.doi.org/10.1017/S0305000910000097

Curtin, S., Campbell, J., & Hufnagle, D. (2012). Mapping novel labels to actions: How the rhythm of words guides infants' learning. *Journal of Experimental Child Psychology*, *112*(2), 127–140. http://dx.doi.org/10.1016/j.jecp.2012.02.007

Curtin, S., Mintz, T. H., & Christiansen, M. H. (2005). Stress changes the representational landscape: Evidence from word segmentation. *Cognition*, *96*(3), 233–262. http://dx.doi.org/10.1016/j.cognition.2004.08.005

Cutler, A. (1986). Forbear is a homophone: Lexical prosody does not constrain lexical access. *Language and Speech, 29*(3), 201–220. http://dx.doi.org/10.1177/002383098602900302

Cutler, A. (2015). Lexical stress in English pronunciation. In M. Reed, & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 106–124). New York, NY: Wiley-Blackwell.

De Vaan, L., Schreuder, R., & Baayen, R. H. (2007). Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon, 2*(1), 1–23. http://dx.doi.org/10.1075/ml.2.1.02vaa

Deacon, S. H., Parrila, R., & Kirby, J. R. (2006). Processing of derived forms in high-functioning dyslexics. *Annals of Dyslexia, 56*(1), 103–128. http://dx.doi.org/10.1007/s11881-006-0005-3

Diependaele, K., Brysbaert, M., & Neri, P. (2012). How noisy is lexical decision? Frontiers in Psychology, 3(348), 1–9. http://dx.doi.org/10.3389/fpsyg.2012.00348

Drewnowski, A., & Healy, A. F. (1982). Phonetic factors in letter detection: A reevaluation. *Memory & Cognition, 10*(2), 145–154. http://dx.doi.org/10.3758/BF03209216

Eddington, D., Treiman, R., & Elzinga, D. (2013). Syllabification of American English: Evidence from a large-scale experiment. Part I. *Journal of Quantitative Linguistics, 20*(1), 45–67. http://dx.doi.org/10.1080/09296174.2012.754601

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18*(2), 116–125. http://dx.doi.org/10.1111/j.1467-9817.1995.tb00077.x

Falkauskas, K., & Kuperman, V. (2015). When experience meets language statistics: Individual variability in processing English compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(6), 1607–1627. http://dx.doi.org/10.1037/xlm0000132

Floccia, C., Nazzi, T., Austin, K., Arreckx, F., & Goslin, J. (2011). Lexical stress and phonetic processing in word learning in 20- to 24-month-old English-learning children. *Developmental Science*, *14*(3), 602–613. http://dx.doi.org/10.1111/j.1467-7687.2010.01006.x

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research, 27*(2), 285–319. http://dx.doi.org/10.1023/A:1023258301588

Ford, M. A., Davis, M. H., & Marslen-Wilson, W. D. (2010). Derivational morphology and base morpheme frequency. *Journal of Memory and Language, 63*(1), 117–130. http://dx.doi.org/10.1016/j.jml.2009.01.003

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*(2), 126–152. http://dx.doi.org/10.1177/002383095800100207

Fudge, E. (1984). *English word-stress*. London: George Allen and Unwin. http://dx.doi.org/10.4324/9781315674810

Goldman, H. B., & Healy, A. F. (1985). Detection errors in a task with articulatory suppression: Phonological receding and reading. *Memory & Cognition, 13*(5), 463–468. http://dx.doi.org/10.3758/BF03198459

Goodman, I., Libenson, A., & Wade-Woolley, L. (2010). Sensitivity to linguistic stress, phonological awareness and early reading ability in preschoolers. *Journal of Research in Reading, 33*(2), 113–127. http://dx.doi.org/10.1111/j.1467-9817.2009.01423.x

Goswami, U., Gerson, D., & Astruc, L. (2010). Amplitude envelope perception, phonology and prosodic sensitivity in children with developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 23*(8), 995–1019. http://dx.doi.org/10.1007/s11145-009-9186-6

Graf Estes, K., & Bowen, S. (2013). Learning about sounds contributes to learning about words: Effects of prosody and phonotactics on infant word learning. *Journal of Experimental Child Psychology*, *114*(3), 405–417. http://dx.doi.org/10.1016/j.jecp.2012.10.002

Greenberg, S. N., Healy, A. F., Koriat, A., & Kreiner, H. (2004). The GO model: A reconsideration of the role of structural units in guiding and organizing text on line. *Psychonomic Bulletin & Review, 11*(3), 428–433. http://dx.doi.org/10.3758/BF03196590

Grimani, A. & Protopapas, A. (2016). Derivational suffixes as cues to stress position in reading Greek. *Journal of Research in Reading.* Advance online publication. http://dx.doi.org/10.1111/1467-9817.12092

Guérard, K., Saint-Aubin, J., Poirier, M., & Demetriou, C. (2012). Assessing the influence of letter position in reading normal and transposed texts using a letter detection task. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Exprimentale, 66*(4), 227–238. http://dx.doi.org/10.1037/a0028494

Gutiérrez-Palma, N. & Palma-Reyes, A. (2008). On the use of lexical stress in reading Spanish. *Reading and Writing: An Interdisciplinary Journal, 21*(6), 645–660. http://dx.doi.org/10.1007/s11145-007-9082-x

Halderman, L. K., Ashby, J., & Perfetti, C. A. (2012). Phonology: An early and integral role in identifying words. In J. S. Adelman (Ed.), *Visual word recognition, Volume I: Models and methods, orthography and phonology*. New York, NY: Psychology Press.

Halle, M., & Vergnaud, J. (1987). *An essay on stress.* Cambridge, MA: The MIT Press.

Harris, L. N., & Perfetti, C. A. (2016). Lexical stress and linguistic predictability influence proofreading behavior. *Frontiers in Psychology, 7*(96), 1–11. http://dx.doi.org/10.3389/fpsyg.2016.00096

Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo results in methodological research: The one- and two-factor fixed effects ANOVA cases. *Journal of Educational Statistics, 17*(4), 315–339. http://dx.doi.org/10.3102/10769986017004315

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago, IL: University of Chicago Press.

Hayes, H., Treiman, R., & Kessler, B. (2006). Children use vowels to help them spell consonants. *Journal of Experimental Child Psychology, 94*(1), 27–42. http://dx.doi.org/10.1016/j.jecp.2005.11.001

Holliman, A. J., Gutiérrez-Palma, N., Critten, S., Wood, C., Cunnane, H., & Pillinger, C. (2017). Examining the independent contribution of prosodic sensitivity to word reading and spelling in early readers. *Reading and Writing: An Interdisciplinary Journal, 30*(3), 509–521. http://dx.doi.org/10.1007/s11145-016-9687-z

Houghton, G., & Zorzi, M. (2003). Normal and impaired spelling in a connectionist dual-route architecture. *Cognitive Neuropsychology, 20*(2), 115–162. http://dx.doi.org/10.1080/02643290242000871

Jarmulowicz, L. (2006). School-aged children's phonological production of derived English words. *Journal of Speech, Language, and Hearing Research, 49*(2), 294–308. http://dx.doi.org/10.1044/1092-4388(2006/024)

Jarmulowicz, L., Hay, S. E., Taran, V. L., & Ethington, C. A. (2008). Fitting derivational morphophonology into a developmental model of reading. *Reading and Writing: An Interdisciplinary Journal, 21*(3), 275–297. http://dx.doi.org/10.1007/s11145-007-9073-y

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434–446. http://dx.doi.org/10.1016/j.jml.2007.11.007

Jarvis, B. (2008a). *DirectRT* (Version 2008.1.0.13) [Computer Software]. New York, NY: Empirisoft Corporation.

Jarvis, B. (2008b). *DirectRT* (Version 2008.2.103.1115) [Computer Software]. New York, NY: Empirisoft Corporation.

Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York, NY: Routledge.

Jouravlev, O., & Lupker, S. J. (2015). Lexical stress assignment as a problem of probabilistic inference. *Psychonomic Bulletin & Review, 22*(5), 1174–1192. http://dx.doi.org/10.3758/s13423-015-0802-y

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development*, *64*(3), 675–687. http://dx.doi.org/10.1111/j.1467-8624.1993.tb02935.x

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, *39*(3–4), 159–207. http://dx.doi.org/10.1006/cogp.1999.0716

Kearns, D. M., Steacy, L. M., Compton, D. L., Gilbert, J. K., Goodwin, A. P., Cho, E., Lindstrom, E. R., & Collins, A. A. (2016). Modeling polymorphemic word recognition: Exploring differences among children with early-emerging and late-emerging word reading difficulty. *Journal of Learning Disabilities, 49*(4), 368–394. http://dx.doi.org/10.1177/0022219414554229

Kelly, M. H. (2004). Word onset patterns and lexical stress in English. *Journal of Memory and Language, 50*(3), 231–244. http://dx.doi.org/10.1016/j.jml.2003.12.002

Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance, 14*(3), 389–403. http://dx.doi.org/10.1037/0096-1523.14.3.389

Kelly, M. H., Morris, J., & Verrekia, L. (1998). Orthographic cues to lexical stress: Effects on naming and lexical decision. *Memory & Cognition, 26*(4), 822–832. http://dx.doi.org/10.3758/BF03211401

Kessler, B., & Treiman, R. (2001). Relationship between sounds and letters in English monosyllables. *Journal of Memory and Language, 44*(4), 592–617. http://dx.doi.org/10.1006/jmla.2000.2745

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language, 47*(1), 145–171. http://dx.doi.org/10.1006/jmla.2001.2835

Klein, R. M., & Saint-Aubin, J. (2016). What a simple letter-detection task can tell us about cognitive processes in reading. *Current Directions in Psychological Science, 25*(6), 417–424. http://dx.doi.org/10.1177/0963721416661173

Koriat, A., & Greenberg, S. N. (1991). Syntactic control of letter detection: Evidence from English and Hebrew nonwords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(6), 1035–1050. http://dx.doi.org/10.1037/0278-7393.17.6.1035

Kriukova, O., & Mani, N. (2016). The strong, the weak, and the first: The impact of phonological stress on processing of orthographic errors in silent reading. *Brain Research, 1636*, 208–218. http://dx.doi.org/10.1016/j.brainres.2016.01.003

Kuperman, V., Drieghe, D., Keuleers, E., & Brysbaert, M. (2013). How strongly do word reading times and lexical decision times correlate? Combining data from eye movement corpora and megastudies. *The Quarterly Journal of Experimental Psychology, 66*(3), 563–580. http://dx.doi.org/10.1080/17470218.2012.658820

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance, 35*(3), 876–895. http://dx.doi.org/10.1037/a0013484

Ladefoged, P., & Johnson, K. (2010). *A course in phonetics* (6th ed.). Boston, MA: Wadsworth, Cenage Learning.

Ling, L. E., Grabe, E., & Nolan, F. (2000). Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech, 43*(4), 377–401. http://dx.doi.org/10.1177/00238309000430040301

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*, 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Marchand, Y., Adsett, C. R., & Damper, R. I. (2009). Automatic syllabification in English: A comparison of different algorithms. *Language and Speech, 52*(1), 1–27. http://dx.doi.org/10.1177/0023830908099881

Mazzoni, D., & Dannenberg, R. (2013). *Audacity* (Version 2.0.3) [Computer Software]. Retrieved from http://audacity.sourceforge.net/

Monaghan, P., Arciuli, J., & Ševa, N. (2016). Cross-linguistic evidence for probabilistic orthographic cues to lexical stress. In J. Thomson, & L. Jarmulowicz (Eds.), *Linguistic rhythm and literacy* (pp. 215–236). Amsterdam/Philadelphia, PA: John Benjamins Publishing Company.

Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: Item response theory analysis of the Author Recognition Test. *Behavior Research Methods, 47*(4), 1095–1109. http://dx.doi.org/10.3758/s13428-014-0534-3

Moravcsik, J. E., & Healy, A. F. (1995). Effect of meaning on letter detection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 82–95. http://dx.doi.org/10.1037/0278-7393.21.1.82

Mousikou, P., Sadat, J., Lucas, R., & Rastle, K. (2017). Moving beyond the monosyllable in models of skilled reading: Mega-study of disyllabic nonword reading. *Journal of Memory and Language, 93*, 169–192. http://dx.doi.org/10.1016/j.jml.2016.09.003

Mundy, I. R., & Carroll, J. M. (2013). Spelling–stress regularity effects are intact in developmental dyslexia. *The Quarterly Journal of Experimental Psychology, 66*(4), 816–828. http://dx.doi.org/10.1080/17470218.2012.719530

Napolitano, D., Sheehan, K., & Mundkowsky, R. (2015). Online readability and text complexity analysis with TextEvaluator. *Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT),* 96–100.

Oller, D. K. (1973). The effect of position in utterance on speech segment duration in English. *The Journal of the Acoustical Society of America, 54*(5), 1235–1247. http://dx.doi.org/10.1121/1.1914393

Paizi, D., Zoccolotti, P., & Burani, C. (2011). Lexical stress assignment in Italian developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 24*(4), 443–461. http://dx.doi.org/10.1007/s11145-010-9236-0

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10*(5), 233–238. http://dx.doi.org/10.1016/j.tics.2006.03.006

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cognitive Psychology, 61*(2), 106–151. http://dx.doi.org/10.1016/j.cogpsych.2010.04.001

Perry, C., Ziegler, J. C., & Zorzi, M. (2014). CDP++.Italian: Modelling sublexical and supralexical inconsistency in a shallow orthography. *PLoS ONE, 9*(4), 1–13. http://dx.doi.org/10.1371/journal.pone.0094291

Protopapas, A., & Gerakaki, S. (2009). Development of processing stress diacritics in reading Greek. *Scientific Studies of Reading, 13*(6), 453–483. http://dx.doi.org/10.1080/10888430903034788

Protopapas, A., Panagaki, E., Andrikopoulou, A., Gutiérrez Palma, N., & Arvaniti, A. (2016). Priming stress patterns in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 42*(11), 1739–1760. http://dx.doi.org/10.1037/xhp0000259

R Core Team. (2017). *R: A language and environment for statistical computing* (Version 3.3.3) [Computer Software]. Vienna, Austria: R Foundation for Statistical Computing.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*(3), 265–292. http://dx.doi.org/10.1016/S0010-0277(99)00058-X

Rastle, K., & Coltheart, M. (2000). Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. *Journal of Memory and Language, 42*(3), 342–364. http://dx.doi.org/10.1006/jmla.1999.2687

Roy-Charland, A., & Saint-Aubin, J. (2006). The interaction of word frequency and word class: A test of the GO model's account of the missing-letter effect. *The Quarterly Journal of Experimental Psychology, 59*(1), 38–45. http://dx.doi.org/10.1080/17470210500269428

Roy-Charland, A., Saint-Aubin, J., Klein, R. M., & Lawrence, M. (2007). Eye movements as direct tests of the GO model for the missing-letter effect. *Attention, Perception, & Psychophysics, 69*(3), 324–337. http://dx.doi.org/10.3758/BF03193753

Saint-Aubin, J., Klein, R. M., & Landry, T. (2005). Age changes in the missing-letter effect revisited. *Journal of Experimental Child Psychology, 91*(2), 158–182. http://dx.doi.org/10.1016/j.jecp.2005.01.007

Saint-Aubin, J., & Poirier, M. (1997). The influence of word function in the missing-letter effect: Further evidence from French. *Memory & Cognition, 25*(5), 667–676. http://dx.doi.org/10.3758/BF03211308

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology, 6*(4), 147–151. http://dx.doi.org/10.1027/1614-2241/a000016

Schwanenflugel, P. J., & Knapp, N. F. (2015). *The psychology of reading: Theory and applications*. New York, NY: Guilford Publications.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523–568. http://dx.doi.org/10.1037/0033-295X.96.4.523

Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry, 11*(3), 563–605.

Ševa, N., Monaghan, P., & Arciuli, J. (2009). Stressing what is important: Orthographic cues and lexical stress assignment. *Journal of Neurolinguistics, 22*(3), 237–249. http://dx.doi.org/10.1016/j.jneuroling.2008.09.002

Sheehan, K. M., Kostin, I., Napolitano, D., & Flor, M. (2014). The TextEvaluator tool: Helping teachers and test developers select texts for use in instruction and assessment. *The Elementary School Journal, 115*(2), 184–209. http://dx.doi.org/10.1086/678294

Sibley, D. E., Kello, C. T., & Seidenberg, M. S. (2010). Learning orthographic and phonological representations in models of monosyllabic and bisyllabic naming. *European Journal of Cognitive Psychology, 22*(5), 650–668. http://dx.doi.org/10.1080/09541440903080583

Small, L. H., Simon, S. D., & Goldberg, J. S. (1988). Lexical stress and lexical access: Homographs versus nonhomographs. *Attention, Perception, & Psychophysics, 44*(3), 272–280. http://dx.doi.org/10.3758/BF03206295

Spinelli, G., Sulpizio, S., Primativo, S., & Burani, C. (2016). Stress in context: Morpho-syntactic properties affect lexical stress assignment in reading aloud. *Frontiers in Psychology, 7*(942), 1–11. http://dx.doi.org/10.3389/fpsyg.2016.00942

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*(4), 402–433. http://dx.doi.org/10.2307/747605

Stone, G. O., & Van Orden, G. C. (1993). Strategic control of processing in word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 19*(4), 744–774. http://dx.doi.org/10.1037/0096-1523.19.4.744

Sulpizio, S., Arduino, L. S., Paizi, D., & Burani, C. (2013). Stress assignment in reading Italian polysyllabic pseudowords. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(1), 51–68. http://dx.doi.org/10.1037/a0028472

Sulpizio, S., & Colombo, L. (2013). Lexical stress, frequency, and stress neighbourhood effects in the early stages of Italian reading development. *The Quarterly Journal of Experimental Psychology, 66*(10), 2073–2084. http://dx.doi.org/10.1080/17470218.2013.785577

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education.

Thiessen, E. D., & Erickson, L. C. (2013). Beyond word segmentation: A two-process account of statistical learning. *Current Directions in Psychological Science, 22*(3), 239–243. http://dx.doi.org/10.1177/0963721413476035

Thomson, J., & Jarmulowicz, L. (Eds.). (2016). *Linguistic rhythm and literacy*. Amsterdam/Philadelphia, PA: John Benjamins Publishing Company. http://dx.doi.org/10.1075/tilar.17

Treiman, R., & Kessler, B. (2006). Spelling as statistical learning: Using consonantal context to spell vowels. *Journal of Educational Psychology, 98*(3), 642–652. http://dx.doi.org/10.1037/0022-0663.98.3.642

Treiman, R., Kessler, B., & Bick, S. (2002). Context sensitivity in the spelling of English vowels. *Journal of Memory and Language, 47*(3), 448–468. http://dx.doi.org/10.1016/S0749-596X(02)00010-4 van Rees, L. J., Ballard, K. J., McCabe, P., Macdonald-D'Silva, A. G., & Arciuli, J. (2012). Training production of lexical stress in typically developing children using orthographically biased stimuli and principles of motor learning. *American Journal of Speech-Language Pathology, 21*(3), 197–206. http://dx.doi.org/10.1044/1058-0360(2012/11-0008)

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology, 47*(3), 631–650. http://dx.doi.org/10.1080/14640749408401131

Venezky, R. L. (1999). *The American way of spelling: The structure and origins of American English orthography*. New York, NY: Guilford Press.

Verrekia, L. D. (1996). *Orthographic representations of lexical stress in English* (Doctoral dissertation). Retrieved from http://repository.upenn.edu/dissertations/AAI9636229

Wade-Woolley, L., & Heggie, L. (2015). Implicit knowledge of word stress and derivational morphology guides skilled readers' decoding of multisyllabic words. *Scientific Studies of Reading, 19*(1), 21–30. http://dx.doi.org/10.1080/10888438.2014.947647

Wade-Woolley, L. & Heggie, L. (2016). The contributions of prosodic and phonological awareness to reading: A review. In J. Thomson, & L. Jarmulowicz (Eds.), *Linguistic rhythm and literacy* (pp. 3–23). Amsterdam/Philadelphia, PA: John Benjamins Publishing Company.

Wade-Woolley, L., & Wood, C. (2006). Prosodic sensitivity and reading development. *Journal of Research in Reading, 29*(3), 253–257. http://dx.doi.org/10.1111/j.1467-9817.2006.00306.x

Wang, M., & Arciuli, J. (2015). Introduction to the special issue. Phonology beyond phonemes: Contributions of suprasegmental information to reading. *Scientific Studies of Reading, 19*(1), 1–4. http://dx.doi.org/10.1080/10888438.2014.976790

Wicherts, J. M., Veldkamp, C. L., Augusteijn, H. E., Bakker, M., van Aert, R. C., & van Assen, M. A. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid *p*-hacking. *Frontiers in Psychology, 7*(1832), 1–12. http://dx.doi.org/10.3389/fpsyg.2016.01832

Wood, C. (2006). Metrical stress sensitivity in young children and its relationship to phonological awareness and reading. *Journal of Research in Reading, 29*(3), 270–287. http://dx.doi.org/10.1111/j.1467-9817.2006.00308.x

Yang, J. & Zevin, J. (2014). The impact of task demand on visual word recognition. Neuroscience, 272, 102–115. http://dx.doi.org/10.1016/j.neuroscience.2014.04.044

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language, 60*(4), 502–529. http://dx.doi.org/10.1016/j.jml.2009.02.001

Yap, M. J., Balota, D. A., Cortese, M. J., & Watson, J. M. (2006). Single- versus dual-process models of lexical decision performance: Insights from response time distributional analysis. *Journal of Experimental Psychology: Human Perception and Performance, 32*(6), 1324–1344. http://dx.doi.org/10.1037/0096-1523.32.6.1324

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*(5), 971–979. http://dx.doi.org/10.3758/PBR.15.5.971

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Sciences Associates.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*(1), 3–29. http://dx.doi.org/10.1037/0033-2909.131.1.3

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for *–ough* and the spelling for /u/? A database for computing feedforward and feedback consistency in English. *Behavior Research Methods, Instruments & Computers, 29*(4), 600–618. http://dx.doi.org/10.3758/BF03210615