University of Southampton Research Repository

UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENVIRONMENTAL AND LIFE SCIENCES

Psychology

Becoming a Skilled Reader: The Development of Parafoveal Pre-processing

by

Sara Victoria Milledge

ORCID ID: 0000-0002-0158-0380

Thesis for the degree of Doctor of Philosophy

April 2021

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENVIRONMENTAL AND LIFE SCIENCES

Psychology

Thesis for the degree of Doctor of Philosophy

BECOMING A SKILLED READER: THE DEVELOPMENT OF PARAFOVEAL PRE-PROCESSING

Sara Victoria Milledge

In contrast to the large body of research that has examined parafoveal pre-processing in skilled adult readers, very little research has examined such processing in beginner child readers. Several novel aspects of parafoveal pre-processing within child readers, in comparison to adult readers, were examined within this thesis; enabling further insight into what information children extract from an upcoming word, and whether there is evidence of developmental change within this processing. Research has shown that adult readers continue to use phonological information to facilitate their lexical identification, counter to current theories of reading development. There appears, rather, to be a developmental change in phonological processing as beginner readers progress to be skilled readers; such that phonology can be used pre-lexically, facilitating lexical identification. This pre-lexical processing of phonology by adults is dependent on orthography. Consequently, prior to examining whether children, like adults, extract phonological information from preview, it was necessary to examine child readers' extraction of orthographic information from preview in my first experiment. Within this experiment it was shown that, firstly, children were sensitive to a word's entire orthographic form in preview, and, secondly, for both adult and child readers, the external letters of an upcoming word were more facilitative to their lexical identification than the internal letters. Moreover, substituting the first letter in preview caused disruption to both adults' and children's processing (first-letter bias). The children's parafoveal pre-processing of orthography was also found to be slower in comparison to that of the adults. My second experiment directly examined whether children process phonological information from an upcoming word, and the extent to which this is affected by orthography. Both adult and child readers benefitted from phonological information being present in preview, though this benefit was modulated by orthography. As such, benefits from phonology were only present within orthographically similar stimuli. Also, the results, again, suggest that both adults and children display a first-letter bias. My third experiment examined this first-letter bias; determining whether it is driven by orthography or phonology. Within both adults' and children's first-pass reading, the first-letter bias was primarily driven by orthography. Again, there was evidence of children's extraction of orthographic information from preview being slower relative to that of the adults. Overall, 8- to 9-year-old child readers' extraction of information from preview was broadly comparable to that of skilled adult readers. There was also evidence found of developmental change in the time course of parafoveal pre-processing of orthography.

Table of Contents

Chapter 5 The Importance of the First Letter in Children's Parafoveal Pre-processing in English: Is it Phonologically or Orthographically Driven? 124

List of Tables

List of Figures

Academic Thesis: Declaration Of Authorship

I, Sara V. Milledge, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Becoming a Skilled Reader: The Development of Parafoveal Pre-processing

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

Milledge, S. V., & Blythe, H. I. (2019). The changing role of phonology in reading development. *Vision, 3*(2), 23. https://doi.org/10.3390/vision3020023 (Chapter 2)

Milledge, S. V., Blythe H. I., & Liversedge, S. P. (2020). Parafoveal pre-processing in children reading English: The importance of external letters. *Psychonomic Bulletin & Review, 28*, 197-208. https://doi.org/10.3758/s13423-020-01806-8 (Chapter 3)

Signed: ....

Date: 20/04/21

Acknowledgements

First and foremost, I would like to thank my external supervisors, Dr Hazel Blythe and Professor Simon Liversedge. The work presented in this thesis would not have been possible without the help, support, and guidance they provided throughout the entirety of my studentship. I would also like to thank my internal supervisors, Dr Denis Drieghe and Dr Hayward Godwin, who were always happy to talk stats with me and provide support when needed.

My sincere thanks also goes to the schools and children who participated in my research (as well as to the parents for allowing their children to participate), and to Robert Chan Seem, David Keep, and Charlotte Lee who endured many early mornings, and the extremes of weather, when accompanying me to collect data.

I would also like to thank, generally, my friends and colleagues within the University of Southampton, past and present, who have helped me both academically and personally throughout my studies. There are too many people to name but their friendship, support, feedback, and encouragement have been invaluable.

I also owe thanks to the ESRC (SCDTP) for funding my M.Sc. and Ph.D.

Finally, I would like to thank my family for their unwavering support. They have helped me in more ways than I can express but, ultimately, have got me to the end of this journey with my sanity mainly intact.

Chapter 1 Introduction

1.1 Reading

Reading is an essential skill in modern societies, as so much information is conveyed by the written word. Poor reading skills can impede an individual's successful functioning within society, both socially and professionally, in terms of cognitive development, academic success, employability, and social and economic welfare. It is vital then to understand how this skill, so crucial to successful functioning within society, develops.

Reading is a complex cognitive process that requires the learning of associations between sequences of printed letters- the visual form of words (orthography)- and their associated speech sounds (phonology), and meanings (semantics). Learning to decode the printed letter string is a critical aspect of word identification which forms the basis of text comprehension- the aim of reading. Readers need to learn how to use the conventional forms of printed language to obtain meaning (Castles et al., 2018; Rayner et al., 2001).

In this thesis, lexical identification was examined within beginner child readers of English compared to skilled adult readers during naturalistic silent sentence reading. Specifically, there is a focus on the roles that orthography and phonology play in facilitating lexical identification of an upcoming word (*n*+1). In the general introduction, within this chapter, a broad overview is provided of the insights that have been gained into both adults' and children's lexical identification processes using the methodology of eye-tracking, indexing eye movement behaviour, during sentence reading (with the basic components of eye movement behaviour outlined and described). Within this chapter, the roles orthography and phonology play regarding processing of word *n*+1 will only very briefly be addressed, due to in-depth discussions present within the introductions of the experimental chapters (Chapters 3, 4, and 5).

Before addressing eye movement behaviour within adult and child readers, a theoretical framework will first be introduced that accounts for the importance of the roles orthography and phonology play in lexical identification: the lexical quality hypothesis (Perfetti, 2007; Perfetti & Hart, 2001, 2002). Although this framework was formulated on the basis of research with adult readers, explaining variation in reading skill (Perfetti & Hart, 2001, 2002), it has also been extended to child readers (Perfetti, 2007).

1.1.1 Lexical Quality Hypothesis

The lexical quality hypothesis (e.g., Perfetti, 2007; Perfetti & Hart, 2001, 2002) emphasises the role of word knowledge in reading skill; that is, the extent to which a reader has access to a lexical entry with representations of its orthography, phonology, and semantics. The quality of lexical representations varies, dependent on the amount of information a reader has for a given word in their lexicon. A high quality lexical representation involves orthographic, phonological, and semantic information being stored with precision (e.g., recognising that *meet* is not the same as *meat*, despite these two words' shared phonology and substantial orthographic overlap). High quality lexical representations are activated in a more synchronous manner, resulting in fast, skilled reading. In contrast, a low quality lexical representation will have at least one of the three key constituents missing or underspecified (e.g., if a word's spelling is unknown or if there is some uncertainty as to its spelling), leading to slower, more effortful, word identification. As such, according to this theory, variability in the quality of lexical representations results in both age and individual differences amongst readers. The more high quality lexical representations a reader has, gained through reading experience, the greater their ability to rapidly lexically identify a word, resulting in fast (skilled) reading. The quality of a word's representations drives the speed with which lexical processing can be completed.

Indeed, it has been theorised that fine-tuning of orthographic processing is key to the development of skilled reading (lexical tuning hypothesis; Castles et al., 1999, 2007), with empirical support having been found for this hypothesis (e.g., Castles et al., 1999, 2007; Kezilas et al., 2017; Polse & Reilly, 2015). According to this hypothesis, specifications of spellings (i.e., orthographic lexical representations) follow a developmental trajectory with regard to their precision. The representations need to develop from being less precise to more precise as a reader's lexicon grows, in order to discriminate between similar words (e.g., *meet* vs. *meat*, with their very different meanings). This development in the precision (i.e., the quality) of orthographic lexical representations facilitates fast lexical identification.

Eye-tracking is a research technique that has enabled researchers to gain insight into visual word recognition processes, including orthographic and phonological processing. It also provides insight into the potential quality of readers' lexical representations during natural sentence reading.

1.2 Eye Movements during Reading

Recordings of eye movement behaviour provide a highly detailed and sensitive index of the moment-to-moment operations involved in the reading process (e.g., Rayner, 1998, 2009; Starr & Rayner, 2001). As such, the use of eye-tracking as a research method affords researchers insight into the psychological processes underlying *where* and *when* the reader moves their eyes, two key decisions for a reader (see Rayner, 1998, 2009 for reviews). It has been found that a number of complex cognitive processing mechanisms underlie these behaviours, and have been modelled in several sophisticated computational models of adult reading (e.g., SWIFT; Engbert et al., 2005; E-Z Reader; Reichle et al., 1998; OB1-reader; Snell et al., 2018).

Within this chapter, the main focus will be on *when* the reader moves their eyes. In the following sections, the basic characteristics of eye movements during reading will be described, firstly considering eye movement behaviour in skilled adult readers and, then, in beginner child readers.

1.2.1 Eye Movement Behaviour in Adult Readers

The two most basic components of eye movements during reading are saccades, that is the rapid movements of the eye themselves, and fixations, wherein there are brief pauses between saccades and the eyes remain fairly motionless and information is visually encoded. Visual input is suppressed during saccades due to the high velocity of the eye movement (Matin, 1974), as high as 500º per second (Rayner, 1998). A saccade's function is to bring new information to the point of fixation, or to return, as a regression, either to a place within the currently fixated word or to a previous word in a sentence (Rayner, 1998, 2009). In adults, fixations during silent reading typically last between 225 and 250 ms (Rayner, 2009), saccades take around 175-200 ms to plan and initiate (Becker & Jürgens, 1979), and saccade duration (the amount of time that it takes to move the eyes) for a 2º saccade, which is typical of reading, normally takes approximately 30 ms (Rayner, 1978). This pattern of eye movements is necessary due to the anatomical constraints of the retina, and its consequent limitations in visual acuity around the point of fixation. The point of highest visual acuity occurs in the fovea, which only occupies around 2º of central vision. Visual acuity continuously decreases as a function of retinal eccentricity, with the parafovea extending 5º outwards on either side of fixation (foveal vision), followed by the periphery, where little visual information can be obtained due to its poor visual acuity (Rayner, 1998, 2009). Thus, readers must make these eye movements in order to place in foveal vision the part of the stimulus they want to see clearly. Both foveal and parafoveal pre-processing are important in skilled reading, as will be discussed later.

The majority of saccades, when reading in English, are made from left-to-right, thereby following the sequential order of words in a given sentence. The mean saccade amplitude is seven to nine letter spaces, and is not modulated by viewing distance (Morrison & Rayner, 1981); however, it is dependent on the processing demands of the text itself, e.g., saccades are shorter if the text is more difficult (Rayner, 1998, 2009). A small proportion of saccades (10-15%) are also made that move backwards in the text (regressions), which can be made either to previous words in the sentence or within the currently fixated word (Rayner, 1998, 2009; Vitu & McConkie, 2000). As text becomes more lexically, semantically, or syntactically difficult or ambiguous (e.g., Frazier & Rayner, 1982), fixation durations and regression rates increase (Rayner, 1998), indicating the importance of the role cognitive processing plays in determining eye movement behaviour during reading.

1.2.1.1 Foveal Processing

Whilst *where* the eyes move appears to be largely determined by low-level visual processes, by variables like word length and space information (e.g., Inhoff et al., 2003; Morris et al., 1990; O'Regan, 1979, 1980; Rayner, 1979), the question of *when* the eyes move is primarily governed by higher order linguistic processes (i.e., lexical processing). For example, a benchmark finding within eye movements during reading is the frequency effect: high frequency words (words often encountered in print) are fixated for a shorter amount of time than low frequency words (words encountered less often in print; e.g., Henderson & Ferreira, 1990; Just & Carpenter, 1980; Liversedge et al., 2004; Rayner & Fischer, 1996), indicating the ease with which lexical access and, thus, lexical identification can occur (Inhoff & Rayner, 1986). Research has found that the frequency effect is evident even when the text disappeared during fixation, after a certain amount of time elapsed (disappearing text paradigm; Liversedge et al., 2004; Rayner et al., 2003; Rayner, Liversedge, et al., 2006). Despite no visual (low-level) information being available to the reader after the word disappeared, participants displayed longer fixations on the subsequent blank space where a low frequency word had been present compared to when a high frequency word had been present. For low frequency words, relative to high frequency words, participants' lexical processing was slower, indicating difficulty in lexical identification. Consequently, they continued to fixate (longer) where a low frequency word had been, as they were still attempting to lexically identify it. Clearly, the difficulty associated with lexically processing a word determines how long readers fixate that given word (or indeed the space where the word was present). In addition, although referring to the question of *where* readers move their eyes, low frequency words are less likely to be skipped than high frequency words (e.g., Brysbaert et al., 2005), demonstrating again that the ease with which a reader can process a word affects their eye movement behaviour. Overall, research supports the notion that there is a tight link between the difficulty of the word being (or to be) lexically processed and readers' fixation behaviour.

Similarly, predictability, given the prior sentence context, has been shown to exert an influence on fixations durations (*when*) and skipping (*where*): more predictable words are fixated for less time, and are more likely to be skipped, than less predictable words (e.g., Drieghe et al., 2005; Ehrlich & Rayner, 1981; Rayner et al., 2011; Rayner & Well, 1996). Also, Age-of-Acquisition (the age at which a reader first encounters a word; e.g., Juhasz & Rayner, 2003, 2006) and word plausibility within a sentence context (e.g., Rayner et al., 2004) have been found to influence reading times on a word. Again, research suggests that lexical processing is the main determinant of *when* adult readers move their eyes, with post-lexical processing also playing a role.

In addition to these variables, there are a number of orthographic factors that can affect fixation times. These include orthographic familiarity of a word (a measure of how commonly the letters within that word tend to appear together; e.g., White, 2008), and the size of a word's orthographic neighbourhood (the number of other words that can be made from one word by changing a single letter; e.g., Pollatsek et al., 1999). Additionally, research has shown that adult readers experience a cost to their processing when letters are transposed (their positions are switched) compared to when words are correctly presented (e.g., Johnson & Eisler, 2012; Rayner, White, et al., 2006; White et al., 2008), with external letter transpositions more harmful than internal letter transpositions to lexical identification, and, regarding external letter manipulations, word-initial letter transpositions more harmful to processing than word-final letter transpositions (e.g., for the word *problem*, *rpoblem* vs. *problme*; Johnson & Eisler, 2012; White et al., 2008). Consequently, letter position information, especially regarding the external letters of a word (particularly the first letter/s), is very important in regard to facilitating adult readers' foveal processing of a word.

Unsurprisingly, given information about a word's orthographic form necessarily comes first within readers' processing, orthographic processing plays an integral role in the lexical processing of adult readers. Moreover, it would appear that certain letters are more critical to lexical identification than others. Overall, it is clear that factors related to the difficulty of the lexical processing associated with a given word heavily influence the amount of time it is fixated for, and plays a role in determining whether a word is fixated at all.[1] (See Chapter 2, Section 2.5.1, for a discussion of how phonological processing affects fixation durations). Thus, the temporal characteristics of eye movements are clearly tightly linked to ongoing lexical processing during skilled adult reading.

1.2.1.2 Parafoveal Pre-processing

In addition to the cognitive processing that occurs as readers directly fixate a word, readers also engage in parafoveal pre-processing: that is, some information from the next word in the sentence (n+1) in the parafovea can be extracted whilst readers are processing the currently fixated word (n). This pre-processing of information in the parafovea is a vital component of skilled reading, as it facilitates the lexical identification of the

---

[1] It is of note, though, that the linguistic factors of word frequency and predictability have relatively small influences on word skipping in comparison to the visual factor of word length (e.g., Brysbaert et al., 2005; Vitu et al., 1995). This reinforces the idea that *where* the eyes move during reading is primarily determined by low-level visual factors.

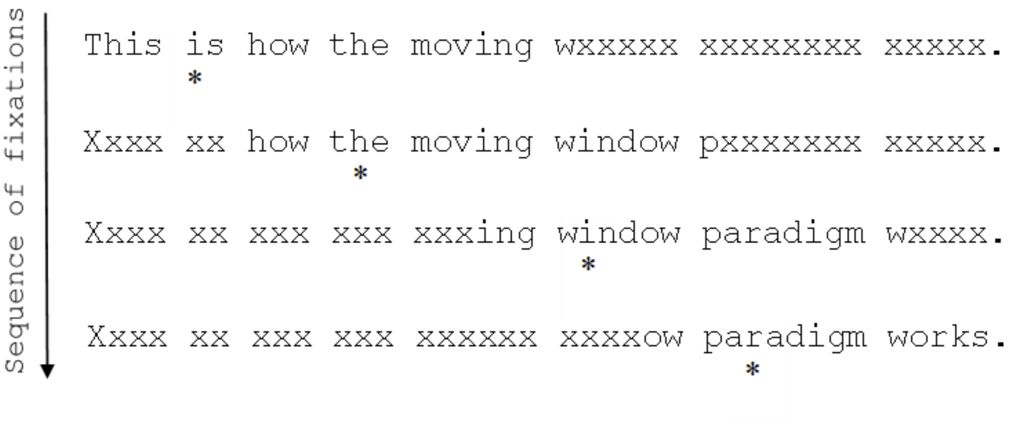
upcoming word when it is subsequently directly fixated (see Schotter et al., 2012 for a review). Indeed, if word *n* disappears or is masked during fixation, there is little disruption to reading; if, however, word *n*+1 is masked or disappears (either at onset of fixation or after 60 ms), adult readers experience considerable disruption to their reading (Rayner, Liversedge, et al., 2006). Thus, the ability to be able to undertake processing of word *n*+1 is critical to the fluency of skilled adult reading.

Two gaze-contingent paradigms, in particular, have been useful in investigating the extent to which readers undertake parafoveal pre-processing during reading, and the type of information that is pre-processed. Firstly, I will discuss the moving window paradigm (McConkie & Rayner, 1975), and the insights it has provided, then the secondary technique of the boundary paradigm (Rayner, 1975) will be discussed.

The moving window paradigm (McConkie & Rayner, 1975) uses a technique wherein the experimenter can manipulate the amount of parafoveal information available to a reader during a given fixation (see Figure 1.1). Using this paradigm, a certain number of characters to the left and to the right of the point of fixation are maintained, creating a window of undisturbed characters that move with the reader's point of fixation, but all characters outside of this range are masked or visually degraded (e.g., replaced with 'x's). By varying the window size (i.e., the number of characters left intact or manipulated) researchers can gain insight into the extent of the perceptual span (the region around the point of fixation from which useful information can be extracted during reading), which enables parafoveal pre-processing. This is done through observing at which window sizes reading is undisturbed, and, equally, the point at which reading speed is disrupted.

Figure 1.1

An example of the moving window paradigm. The reader's fixation location is represented by the asterisk underneath each sentence. In this example, the window size is 4 characters to the left and 14 characters to the right of the point of fixation.



It has been found that the perceptual span extends, for skilled adult readers in an alphabetic language (read from left-to-right), over an asymmetrical area from approximately 3-4 character spaces to the left of fixation, to 14-15 character spaces to the right of fixation; that is reading rate (words per minute) reached asymptote at these window sizes (McConkie & Rayner, 1975). As discussed previously, visual acuity decreases as a function of distance from the fovea, but there is no anatomical explanation as to why this asymmetry is observed in the perceptual span. Linguistic processing appears to drive this effect. It has been found that the properties of a language's writing system affect the asymmetry, and size, of the perceptual span. For example, it has been found that the opposite pattern of asymmetry occurs in readers of Arabic and Hebrew, languages which are read right-to-left, such that the perceptual span extends further to the left than to the right of the point of fixation (Jordan et al., 2014; Pollatsek et al., 1981). Also, in Chinese and Tibetan, where orthographic information is denser than in alphabetic languages, it has been found that the perceptual span is smaller (Inhoff & Liu, 1998; Wang et al., 2020, respectively). In addition, factors like reading ability (e.g., Veldre & Andrews, 2014) and speed (e.g., Ashby et al., 2012; Rayner et al., 2010) have also been found to affect the size of the perceptual span. As such, research suggests that the perceptual span is determined by ongoing processing demands; it is variable in asymmetry and size, dependent on the demands of the language, as well as the processing demands an individual reader faces within their processing.

Secondly, the boundary paradigm (Rayner, 1975) has been used (see Figure 2.2, pp. 46-47). Whilst the moving window paradigm provides an estimate of how much information readers are able to extract during a single fixation, it cannot provide detail as to what exact form of information is extracted. In contrast, the boundary paradigm affords insight into what information and specific characteristics of a word can be parafoveally pre-processed. This paradigm involves an invisible boundary being placed in a sentence, in the space immediately before a target word. A preview letter string is available in the target word's location prior to the reader making a saccade that crosses this invisible boundary. When the reader's eyes cross the boundary, and they subsequently directly fixate the target word, then a display change occurs wherein the preview letter string changes to the correct target word. The use of this paradigm has shown that when a reader is given a correct preview of a target word (i.e., no display change occurs and the target word and preview are identical), their lexical identification is faster than when an incorrect (manipulated) preview is provided. This facilitation to lexical identification is referred to as preview benefit. The fact that adult readers display preview benefit indicates that they extract and begin to process information from the preview letter string prior to fixation, which they subsequently integrate with information obtained on the next fixation, typically made on the target word itself. Given such processing, it is possible to determine the specific types of information readers extract from word *n*+1 to facilitate their lexical identification, by manipulating certain characteristics of the overlap between the preview letter string and the target word that will be directly fixated.

A large body of research has found that skilled adult readers can pre-process information regarding word spacing (e.g., Epelboim et al., 1997; Johnson & Eisler, 2012; Morris et al., 1990; Perea & Acha, 2009; Pollatsek & Rayner, 1982; Rayner et al., 1998; Spragins et al., 1976) and word length (e.g., Inhoff et al., 2003; Juhasz et al., 2008). Moreover, of critical importance to this thesis, adult readers extract orthographic (e.g., Balota et al., 1985; Binder et al., 1999; Johnson et al., 2007; Johnson & Dunne, 2012; Jouravlev & Jared, 2018; McConkie & Zola, 1979; Rayner et al., 1980, 1986, 2014; White et al., 2008) and phonological information (e.g., Ashby & Rayner, 2004; Ashby et al., 2006; Chace et al., 2005; Henderson et al., 1995; Jouravlev & Jared, 2018; Miellet & Sparrow, 2004; Pollatsek et al., 1992; Sereno & Rayner, 2000) from preview. For in-depth discussions regarding parafoveal pre-processing of orthography see the introductions of the experimental chapters (Chapters 3, 4, and 5), and see Chapters 2, 4, and 5 for details of research that has examined phonological processing of word *n*+1.

Whilst these characteristics have been shown to be parafoveally pre-processed by adult readers, it does not appear that semantic information is typically pre-processed in English (e.g., Rayner et al., 1986, 2014). It has, however, been found to be pre-processed in non-alphabetic languages, like Chinese and Korean (e.g., Yan et al., 2009; Yan, Wang, et al., 2019; Yang et al., 2012).

It has been found that the amount of parafoveal pre-processing a reader undertakes can be constrained by the difficulty of processing associated with the fixated word (*n*); this is referred to as foveal load. If word *n* is harder to process, foveal load is high, and less information is pre-processed from word *n*+1. For example, Henderson and Ferreira (Experiment 1; 1990) manipulated the availability of parafoveal word information and foveal load, through a word frequency manipulation. It was found that less processing of word *n*+1 was undertaken when foveal load was high (i.e., when word *n* was a low frequency word). Importantly, this effect cannot be explained simply by the relative amounts of time available to the reader for pre-processing of word *n*+1 (in fact, were pre-processing simply driven by the amount of time available to the reader, then the opposite pattern of effects would be predicted). Specifically, fixations on low frequency words are typically longer in duration (e.g., Rayner, 1998, 2009); a consequence of this is that the reader has a greater amount of time available for parafoveal pre-processing of word *n*+1. In contrast, during a comparatively shorter fixation on a high frequency word, less time is available to the reader for pre-processing of word *n*+1. Contradictory to what this might suggest, Henderson and Ferreira's (1990) data show that word *n*+1 received more pre-processing after the shorter fixation on a high frequency word. This suggests, therefore, that the ease of processing word *n* is a determinant of the extent to which attention can be allocated to word *n*+1 for parafoveal pre-processing.

This finding of foveal load, though, was suggested by Schroyens et al. (1999) to be caused by spillover processing following short preceding fixations. Short fixations on word *n* could lead to the ongoing processing of word *n* 'spilling over' onto word *n*+1, thereby interfering with the extent to which the reader could pre-process word *n*+1. Low frequency words, due to increased difficulty in lexical processing, would be more likely to produce such effects than high frequency words; leading to reduced parafoveal pre-processing for low frequency words with short, but not long, preceding fixations. White et al. (2005), however, found that when word *n* was low frequency this did not affect reading times on word *n*+1, and the duration of the previous fixation played no role in the degree to which parafoveal pre-processing was undertaken. Consequently, White et al. (2005), in addition to replicating the foveal load effect found by Henderson and Ferreira (1990), concluded that this effect was not due to spillover processing.

Overall, the foveal load effect suggests that if more processing resources are necessary for identification of word *n*, then fewer resources are available to start processing word *n*+1, complementing research showing how linguistic processing determines the spatial extent of the perceptual span (e.g., Jordan et al., 2014; Veldre & Andrews, 2014). However, the replicability of the foveal load effect has been questioned (e.g., Drieghe et al., 2005; Findelsberger et al., 2019; Veldre & Andrews, 2018; Zhang et al., 2019).

1.2.1.3 Summary

It is clear that a reader's pattern of eye movements provides insight into the cognitive processes that occur during reading. A word's lexical characteristics, such as its frequency (e.g., Henderson & Ferreira, 1990; Inhoff & Rayner, 1986) and predictability (e.g., Drieghe et al., 2005; Ehrlich & Rayner, 1981; Rayner et al., 2011; Rayner & Well, 1996) influence whether a word is fixated during reading, and, indeed, how long it is fixated for. Fixation durations are an index of the cognitive processes of lexical identification that underlie reading, not just reflecting the visual uptake of text, as has been shown through research using the disappearing text paradigm (Liversedge et al., 2004; Rayner et al., 2003; Rayner, Liversedge, et al., 2006). In addition, findings from the boundary paradigm (Rayner, 1975) have shown that parafoveal pre-processing is facilitative to lexical identification when a word is subsequently directly fixated, and that orthographic and phonological information for this word (*n*+1) is processed prior to it being directly fixated. As such, the ease with which lexical processing is, and can be, undertaken is a key determinant of eye movement behaviour within skilled adult readers, affecting both foveal and parafoveal pre-processing. The following section will discuss the similarities and differences observed between beginner child and skilled adult readers.

1.2.2 Eye Movement Behaviour in Child Readers

Compared to the vast literature on adults' eye movements during reading, relatively few studies have been conducted to investigate the pattern of children's eye movements during reading. This is mainly due to the difficulties associated with tracking children's eye movements and the requirements of the systems used in the past (e.g., the need for bite bars). As a result, however, of technological advances within video-based recording systems, these barriers to conducting eye movement research with children have been removed. This has led to increasingly more research being conducted into children's eye movements during reading (see Blythe & Joseph, 2011 for a review), though there are still clear gaps in knowledge. Nonetheless, this research has led to differences in patterns of eye movements between children and skilled adult readers being identified. Moreover, these age-related changes in eye movement behaviour during reading have been demonstrated to occur predominantly as a consequence of increases in children's lexical processing abilities, as discussed later.

Age-related changes in eye movement behaviour are now well documented: as chronological age increases, fixation durations and sentence reading times decrease, saccadic amplitudes increase, fewer regressions and fixations are made, the probability of word skipping increases, and the probability of making refixations decreases (Blythe et al., 2006, 2009, 2011; Buswell, 1922; Häikiö et al., 2009; Huestegge et al., 2009; Joseph et al., 2009; McConkie et al., 1991; Rayner, 1986; Taylor, 1965). For example, the average fixation duration for 6- to 7-year-old children has been found to be 432 ms (Buswell, 1922; McConkie et al., 1991), with around 200 fixations made per 100 words (McConkie et el., 1991; Taylor, 1965); however, by 11- to 12-years of age, the average fixation duration is around 270 ms (Buswell, 1922; Taylor, 1965), with an average of around 120 fixations per 100 words (Taylor, 1965).

1.2.2.1 Foveal Processing

Research has also found that, similar to adults, there is strong evidence for lexical processing influencing children's eye movements during reading. For example, word frequency exerts a strong influence on children's eye movements, with low frequency words receiving longer fixations than high frequency words (Blythe et al., 2009; Huestegge et al., 2009; Hyönä & Olson, 1995; Joseph et al., 2013; Valle et al., 2013). The finding that lexical characteristics can affect fixation durations during reading, even in children as young as 7-years-old, demonstrates that eye movements during reading are influenced by lexical processes from a young age. Additionally, results from the use of the disappearing text paradigm provide further evidence for the assertion that children's eye movements during reading are under cognitive control. Blythe et al. (Experiment 1; 2009) demonstrated that children, similar to adults, were no more disrupted in terms of their reading speed and comprehension accuracy when words disappeared from a screen after 60 ms, than during normal reading, also displaying marked frequency effects like in adults (e.g., Rayner et al., 2003). In addition, it has been found that phonological processing (see Chapter 2, Section 2.5.2) and letter transpositions (an orthographic manipulation) also affect children's fixation durations. Pagán et al. (2021), within a reading-like task, found that when letters were transposed, compared to a correctly spelled word (e.g., *bandage* vs. *abndage*, *nabdage*, *bnadage*), children's reading times significantly increased. This suggests that both adults' and children's eye movement patterns, especially regarding the decision of *when* the eyes move during reading, reflect the ease with which words can be successfully identified within text.

Moreover, children's eye movement behaviour has also been found to be affected by post-lexical processing. Joseph et al. (2008) found that children's processing was more difficult, shown by longer reading times, for implausible target words (e.g., *Dad used a sword to protect the purple flowers in the garden.*) than plausible target words (e.g., *Dad* *used a fence to protect the purple flowers in the garden.*). This indicates, once again, the important role cognitive processing plays in determining both adults' and children's eye movement behaviour during reading, especially regarding how long words are fixated for in a sentence.

That is not to say though that low-level visual factors do not play a role in determining children's eye movement behaviour during reading. For example, it has been found that word length affects how long children will fixate a word for in a sentence (if they fixate it at all). Children typically make longer fixations on long words than short words (Blythe et al., 2011; Huestegge et al., 2009; Hyönä & Olson, 1995; Joseph et al., 2009), as has been observed in adults (e.g., Rayner et al., 2011), and, like adults, are more likely to skip short words than long words (Joseph et al., 2009). Interestingly, the word length effect (i.e., long vs. short words) has been found to be greater in younger compared to older children (Huestegge et al., 2009). This suggests that not only do younger child readers experience an increased processing load when reading long as compared to short words, but this changes with age as, in contrast to 8- to 9-year-old children, 10- to 11-year-old children (like adults) do not generally require a second visual sample on long words (Blythe et al., 2011). It would appear then that children gradually become faster at encoding visual information from text, allowing lexical processing to begin.

Overall, like skilled adult readers, children's eye movement behaviour during reading is under cognitive control, especially regarding *when* the eyes move. Broadly, children's eye movements are similar to those of adults (i.e., both display word frequency and length effects; e.g., Blythe et al., 2009; Hyönä & Olson, 1995). The question remains, though, as to why there are age-related changes in eye movement behaviour during reading (i.e., why children, in comparison to adults, make longer and more fixations, more regressions, less likely to skip words, etc.; e.g., Blythe et al., 2006, 2009, 2011; Joseph et al., 2009; McConkie et al., 1991).

1.2.2.2 What Drives the Developmental Change in Adults’ and Children’s Eye Movements during Reading?

Although research suggests that there might be some age-related trends of maturation in basic oculomotor control (e.g., in the stability and control of fixations; Aring et al., 2007), they do not appear to be responsible for the developmental change between children’s and adults’ eye movements during reading. For example, children appear to require a comparable amount of time to adults to extract visual information during a fixation (even as fast as 40 ms- Experiment 2; Blythe et al., 2009), and can target their saccades similarly to adults towards the centre of words (Joseph et al., 2009; McConkie et al., 1991). Therefore, the differences between adults’ and children’s oculomotor control and visual processing systems are not responsible for the differences between their eye movements during reading (i.e., how children make longer fixation durations, more regressions and fixations, etc. than skilled adult readers). As such, differences on a cognitive, lexical, processing level, are the primary reason for these differences (i.e., the speed with which skilled adult readers, in comparison to beginner child readers, can achieve lexical identification).

Research confirming the role that lexical processing plays in determining children’s eye movements during reading (e.g., Blythe et al., 2009), is broadly consistent with the lexical quality hypothesis (as discussed in Section 1.1.1). Whilst this hypothesis primarily explains variation between adult readers regarding the quality of their lexical representations and their reading skill (e.g., Perfetti & Hart, 2001, 2002), one study has tested the suggestion that lexical representations are of lower quality within younger readers. If younger readers have lower quality representations of a given word, lexical identification of that word would be more difficult and, thus, slower. As younger readers gain increasingly more reading experience, more lexical representations will be generated as new words are encountered, and, given repeated exposure to words, individual lexical representations become of higher quality (e.g., Perfetti, 2007). This developmental increase in the number of high quality lexical representations could lead to more adult-like eye movement behaviour during reading.

Luke et al. (2015) tested the quality of young teenage readers' lexical representations through a word comprehension and a text comprehension test. Both of these tests required high quality semantic and orthographic representations (i.e., comprehension is not possible without access to meaning, and representations needed to be precise and available simultaneously so that participants were able to access a word's orthographic form in response to semantic cues). It was found that young teenagers with higher lexical quality scores had eye movement patterns more similar to those of adults in a reading task (shorter gaze durations and lower probability of refixations being made). This suggests that eye movements during reading become faster with increased quality of lexical representations. Importantly, the influence of lexical quality was specific to reading: fixation durations in reading diverged from the other tasks they completed as lexical quality increased (there were significantly shorter fixations in reading compared to scene search and pseudoreading). Consequently, the findings suggest that eye movements become increasingly tuned to written language processing as lexical representations become of higher quality due to reading experience.

Given younger child readers' lack of reading experience, their lexical representations would be of lower quality than those of adults (e.g., Perfetti, 2007), and this could help to explain why their eye movements during reading are indicative of more effortful lexical processing. Due to underspecification, and potentially missing constituent information, of their lexical representations, children's lexical processing cannot be as fast as that of skilled adult readers: theoretically, although also shown empirically below, children's and adults' lexical processing differs.

Of the multiple models that have been created to account for adult eye movement behaviour during reading (e.g., SWIFT; Engbert et al., 2005; E-Z Reader; Reichle et al., 1998; OB1-reader; Snell et al., 2018), to date there is only model that has been extended to children- the E-Z Reader. The E-Z Reader model (Reichle et al., 1998) assumes that lexical processing is serial, such that attention is allocated to one word at a time, and that the initiation of a saccade is tightly linked to the lexical processing of the fixated word. The model is composed of three core components: saccadic programming (oculomotor control), lexical processing, and higher language processing. It is the first two of these components that have been focused on in order to provide insight into what drives the differences between adults' and children's eye movement behaviour during reading.

Reichle et al. (2013) and Mancheva et al. (2015) ran simulations using the E-Z Reader model to explain various eye movement phenomena in children. The researchers found that whilst the basic pattern of eye movements exhibited by children (i.e., longer fixations and shorter saccades with more regressions) could not be generated by varying the values of any of the model's parameters that regulate oculomotor control (i.e., the timing and/or accuracy of saccadic programming and/or execution), the basic pattern could be generated by simply reducing the overall rate of lexical processing in children. Thus, slower lexical processing is argued to be the primary determinant of age-related changes between adults' and children's eye movements during reading, rather than differences in oculomotor control.

In sum, the evidence strongly suggests that children's lexical processing is slower than that of skilled adult readers (potentially as a result of children having fewer high quality lexical representations; e.g., Perfetti, 2007), driving the developmental change in their eye movement behaviour. This has been evidenced not only in global measures across whole sentences (e.g., children typically making longer and more fixations, displaying longer sentence reading times, making fewer regressions and refixations, etc.; Blythe et al., 2006, 2009, 2011; Häikiö et al., 2009; Huestegge et al., 2009; Joseph et al., 2009; McConkie et al., 1991), but also in local measures examining lexical identification of a target word within a sentence context. For example, a consistent finding is group differences within reading times, such that children, despite their lexical processing being similar to that of adults (i.e., displaying orthographic and word frequency effects; e.g., Blythe et al., 2009; Pagán et al., 2021), display longer reading times than the adults. This suggests that although typical child readers, as young as 7-years-old, have developed to a level, like adult readers, where their lexical processing of the fixated word is the primary determinant of *when* they move their eyes during reading (e.g., Blythe et al., 2009), their lexical processing is still slower than that of skilled adult readers (e.g., Blythe et al., 2009, 2011, 2015; Joseph et al., 2009, 2013; Mancheva et al., 2015; Pagán et al., 2021; Reichle et al., 2013). To be clear, it seems likely that the time course of lexical processing drives developmental change in eye movement behaviour during reading.

Of note is the fact that manipulations of lexical processing difficulty (e.g., word frequency and predictability) with skilled adult readers typically lead to their eye movements becoming more like those of beginner child readers (i.e., longer fixations, less word skipping, etc.; Rayner, 2009). Consequently, the differences in eye movement behaviour displayed in children compared to adults are generally considered to be a consequence, not the cause, of the reader's processing difficulty. Overall, the observation of similar eye movement behaviour phenomena in child and adult readers (i.e., the frequency effect under disappearing text conditions; Blythe et al., 2009) suggests that the key constant between these two groups is the influence that lexical processing exerts on eye movements during reading.

1.2.2.3 Parafoveal Pre-processing

Parafoveal pre-processing is a core part of fast, skilled adult reading (e.g., Rayner, Liversedge, et al., 2006); yet how this ability develops in, and what information is extracted from word *n*+1 by, beginner child readers of English is largely unknown. Only a small number of studies have investigated parafoveal pre-processing in children using the moving window (McConkie & Rayner, 1975) and boundary (Rayner, 1975) paradigms. These gaps in knowledge started to be addressed in the experiments reported within this thesis; furthering our understanding of how children progress from beginner to skilled reader, by examining novel aspects of parafoveal pre-processing within child readers, and comparing their processing to that of adult readers, using the boundary paradigm.

The amount of letter and word length information that can be extracted from the parafovea, within alphabetic languages, has been shown to undergo developmental change, as the perceptual span increases: whilst 11-year-old children have perceptual spans roughly equivalent to those of adults, 7- to 9-year-old children have smaller perceptual spans of 3-4 letter spaces to the left of fixation and 11 letters to the right (Häikiö et al., 2009; Rayner, 1986; Sperlich et al., 2015), with around one year of reading experience necessary to develop the asymmetry of the span (Rayner, 1986).

Similar to adults, children’s perceptual spans are determined by ongoing processing demands. For example, text difficulty (Rayner, 1986) and reading ability (Yan et al., 2020) have been found to affect the size of the perceptual span; if a child reader experiences increased processing difficulty, be it due to text difficulty or low reading skill, their perceptual span becomes (temporarily) smaller. The size of children’s letter identity spans (the distance from fixation at which readers can access letter-specific information) is also dependent on reading skill and speed (Häikiö et al., 2009). Altogether, the perceptual/letter identity spans of both adults and children are driven by linguistic processing; they are variable in size and asymmetry, dependent on reading skill (e.g., Häikiö et al., 2009; Veldre & Andrews, 2014) and the demands imposed on the reader by the text, be it in terms of processing difficulty (Rayner, 1986) or the language being read (e.g., Pollatsek et al., 1981; Wang et al., 2020).

Moreover, this suggests that even in developing beginner readers, information about upcoming letters and the next word in a sentence can be parafoveally pre-processed during reading. These aforementioned changes that occur in the perceptual/letter identity span, typically with age, do suggest though that children might be more constrained as to their ability to extract information from word *n*+1. This could have contributed to why younger children (8- to 9-year-olds) are more likely to refixate, or require a second visual sample on, long words compared to short words, in comparison to older children (10- to 11-year-olds) and adults (Blythe et al., 2011). Younger children may require more visual samples of a long word in order to encode it sufficiently for normal lexical processing to proceed, due to a lack of parafoveal pre-processing having occurred as a consequence of their more limited perceptual/letter identity spans (e.g., Häikiö et al., 2009; Rayner, 1986; Sperlich et al., 2015).

Again, this research (Blythe et al., 2011) also indicates that younger children's foveal processing is more effortful, shown by multiple fixations being made, than that of older child and skilled adult readers. As such, it would appear that typically developing readers increasingly display faster lexical processing during reading, during direct fixation (also implicated by the age-related changes in eye movement behaviour discussed previously; e.g., Blythe & Joseph, 2011). Conceivably, given typical age-related improvements in lexical processing of word *n*, and reading skill, more attention can be allocated to extracting information from word *n*+1, as evidenced by how age and reading skill increase the perceptual/letter identity span (e.g., Häikiö et al., 2009; Rayner, 1986).

It is very likely that children's reduced rate of lexical processing, relative to that of adults, would also affect their ability to extract information from word *n*+1. There is converging evidence of children's processing being slower than that of adults (e.g., Blythe et al., 2009, 2011, 2015; Joseph et al., 2009, 2013; Mancheva et al., 2015; Pagán et al., 2021; Reichle et al., 2013; Tiffin-Richards & Schroeder, 2015b), with this reduced rate of lexical processing being primarily behind the typical age-related differences between adults' and children's eye movements during reading, rather than differences in oculomotor control (Mancheva et al., 2015; Reichle et al., 2013). The slower speed of children's lexical identification (demonstrated by longer reading times) is indicative of more resource-intensive processing; the more resources that are required for lexical identification of word *n*, then the fewer resources, presumably, are available for processing of word *n*+1.

Of note, research examining foveal processing within both adult and child readers has found that children process the same information as adults within the first fixation on a word (e.g., word frequency; Blythe et al. 2009; Tiffin-Richards & Schroeder, 2015b). However, despite extracting similar information from word *n*, children invariably display longer fixation durations on word *n*, overall, than adults. This is likely to have an effect on the time course with which children are able to start processing word *n*+1, such that their parafoveal pre-processing is likely to be slower than that of adults. Consistent with this notion of developmental change within the time course of parafoveal pre-processing, research has found that children's reading times, compared to those of adults, are longer on word *n*+1 when it is directly fixated (Häikiö et al., 2010; Pagán et al., 2016; Tiffin-Richards & Schroeder, 2015a). This is indicative of children's lexical processing occurring at a slower rate, leading to overall group differences in reading times on word *n*+1.

There is a very small existing body of research that has examined what information child readers are able to extract from word *n*+1 using the boundary paradigm (Rayner, 1975) in alphabetic languages (in Finnish: Häikiö et al., 2010; in German: Marx et al., 2015; Tiffin-Richards & Schroeder, 2015a; and in English: Johnson et al., 2018; Pagán et al., 2016).

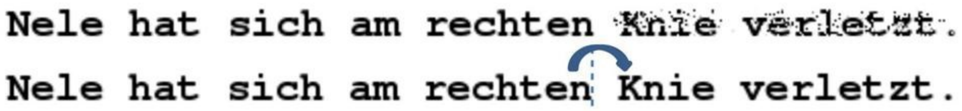
This research has found that children are able to parafoveally pre-process orthographic information (at least partially; Häikiö et al., 2010; Johnson et al., 2018; Marx et al., 2015; Pagán et al., 2016; Tiffin-Richards & Schroeder, 2015a) and are able to pre-process phonology as well (Tiffin-Richards & Schroeder, 2015a). For in-depth discussions regarding parafoveal pre-processing of orthography see the introductions of Chapters 3, 4, and 5, and see Chapter 2 for a discussion of processing of phonology from preview within languages other than English (Section 2.5.2). No research has been conducted to examine parafoveal pre-processing of semantics within alphabetic languages, but semantic information has been shown to be pre-processed by older child readers, 10- to 11-year-olds, in Chinese (Yan, Liu, et al., 2019).

Here, a brief overview is provided as to what is known about child readers' parafoveal pre-processing in English, and the gaps in our knowledge regarding this processing. Only two studies, to date, have been conducted, which have shown that child readers are, at least partially, sensitive to a word's orthographic form in the parafovea, and that the first letter/s of words are important to children's parafoveal pre-processing (Johnson et al., 2018; Pagán et al., 2016; see introductions to Chapters 3, 4, and 5). Similar to past research with adults (e.g., White et al., 2008), children appear to display a bias towards the first letters of a word, with perhaps the first letter playing an especially important role in their ability to lexically identify a word (consistent with non-eye-tracking data; e.g., Marchbanks & Levin, 1965; Rayner & Kaiser, 1975; Williams et al., 1970). Whether both external letters (i.e., first and final), like in adults (e.g., White et al., 2008), are facilitative to lexical identification, though, is unknown. Consequently, it is also unknown whether external letters are more facilitative than internal letters to children's lexical identification processes. Nevertheless, for both child and adult readers of English, the first letter, at least, appears to be a critical lexical access unit for word identification during reading, but why exactly this letter plays such an important role is unknown. Likewise, it is unknown if child readers of English, like skilled adult readers (see Chapters 2 and 4; e.g., Pollatsek et al., 1992), can extract phonological information from word *n*+1. In addition, given the very limited body of research, further research is required in order to determine whether there is indeed evidence of developmental change in the time course over which readers of English extract information from preview.

Before discussion of the research that has examined the modulation of parafoveal pre-processing in children, it is necessary to introduce another paradigm: the incremental boundary paradigm (e.g., Findelsberger et al., 2019; Gagl et al., 2014; Marx et al., 2015, 2016, 2017). This paradigm is similar to the traditional boundary paradigm (Rayner, 1975) but instead of using parafoveal letter masks in preview, researchers manipulate the salience of the parafoveal previews by reducing their visual integrity (degrade them), through the displacement of a certain number of pixels in preview (see Figure 1.2).

Figure 1.2

An example of the incremental boundary paradigm. The top sentence demonstrates how the salience of the preview was manipulated; in this example the amount of displaced pixels is 10%. The bottom sentence shows where the invisible boundary was present (indicated by the dashed line) and how, with the movement of the readers' eyes (shown by the arrow) across this boundary, the degraded preview changed to the full target word (i.e., no pixels displaced).



Nele hat sich am rechten Knie verletzt.

Nele hat sich am rechten Knie verletzt.

(Marx et al., 2016, p. 4)

Using this paradigm, in contrast to previous research conducted with adults using the traditional boundary paradigm (e.g., Henderson & Ferreira, 1990; White et al., 2005), Marx et al. (2017) found that young German readers did not exhibit an effect of foveal load but a substantial spillover effect (i.e., the processing of the first word 'spills over' onto, and affects the processing, of the next word).

Marx et al. (2017) found a similar magnitude of preview benefit across the degraded preview conditions, independent of a foveal load manipulation (whether word *n* was low or high in frequency). The frequency manipulation of word *n*, however, did elicit an effect on gaze durations: if word *n* was low frequency, this resulted in longer gaze durations on word *n*+1 (when it became word *n*). Therefore, it would appear that the young readers were displaying a spillover effect of the frequency of word *n* onto word *n*+1 when it was directly fixated, albeit a late effect emergence (i.e., not found in first fixation duration), and not a modulation to the extent to which they were able to undertake parafoveal pre-processing of word *n*+1 due to foveal load. Marx et al. do concede though that this lack of effect of foveal load, present in adults (e.g., Henderson & Ferreira, 1990) but seemingly absent in children, may be due to other factors (i.e., beginner child readers might not display this effect but skilled adult readers do). The lack of an adult control group does limit their conclusions. Findelsberger et al. (2019), however, have found evidence of foveal load within skilled adult readers when using degraded previews. This tentatively supports the suggestion that beginner child readers might not (yet) display an adult-like effect of foveal load on their ability to extract information from word *n*+1. Lexical processing, though, was clearly still playing a role in determining the children's eye movements during reading: when lexical processing was more difficult (i.e., word *n* was low frequency), this slowed their processing of word *n*+1 when it was directly fixated.

Again, the role lexical processing plays in helping to determine eye movement behaviour during reading is clear. Not only does processing difficulty affect foveal processing (e.g., Blythe et al., 2009; Liversedge et al., 2004) but also seemingly parafoveal pre-processing in both adults and children. This is apparent through either modulation of the reader's ability to pre-process word *n*+1 (foveal load; e.g., Henderson & Ferreira, 1990), or the reflection of ongoing processing of word *n*, due to its difficulty (spillover effects), on word *n*+1 when it is subsequently directly fixated and becomes word *n* (e.g., Findelsberger et al., 2019; Marx et al., 2017).

It is important to note that the extent to which a reader extracts information during parafoveal pre-processing can be modulated by the type of mask that is used in preview. It has been proposed that the classic boundary paradigm's (Rayner, 1975) use of parafoveal letter masks does not provide an accurate estimate of preview benefit (e.g., Kliegl et al., 2013). A proposed alternative is the incremental boundary paradigm's (e.g., Marx et al., 2015, 2016, 2017) use of degraded previews (as described above; see Figure 1.2).

One piece of research has directly examined whether parafoveal letter masks interfere with children's lexical identification in German. Marx et al. (2015) used valid (identity) previews or parafoveal letter masks ('x's or orthographically similar letter substitutions in preview; e.g., *Xxxx* or *Zwmn* as previews for *Haus*) within the incremental boundary paradigm, so the degradation of the previews was also manipulated. Using this technique, it is proposed that researchers can determine whether a manipulation facilitates or interferes with processing. If decreasing degradation of the previews leads to faster reading times, then the previews facilitated lexical identification; if, however, decreasing degradation results in slower reading times, then the previews have interfered with lexical identification. Marx et al. found that the parafoveal letter masks induced processing costs; the more visually salient and less degraded a preview was, the longer reading times were for the word when it was directly fixated. In contrast, for the valid previews, decreasing degradation led to faster reading times (i.e., lexical identification was facilitated). Interestingly, the data also suggests that increasing degradation of the invalid previews (orthographically similar letter substitutions and 'x' masks) decreased fixation durations, resulting in fixation durations that were more comparable to those displayed on the identity previews (with commensurate degradation). Overall, though, given the processing costs associated with the parafoveal letter masks, this would lead to an overestimation of preview benefit- the extent to which readers benefit from a valid preview relative to an invalid (manipulated) preview.

Whether degraded previews, in comparison to parafoveal letter masks, reduce processing costs within skilled adult reading in English has also been examined, based on an implication from Marx et al.'s (2015) results that degrading invalid masks reduced processing costs. Vasilev et al. (Experiment 1; 2018) compared identity previews (e.g., *garden*), undegraded letter mask previews (e.g., *quvlas* as a preview for *garden*), and the same letter mask previews (e.g., *quvlas*) but degraded, and found that all preview conditions led to preview benefit. Moreover, Vasilev et al. found that, generally, visually degraded previews did not reduce processing costs (fixation durations) compared to undegraded letter mask previews. In fact, previews degraded by 20% led to longer fixation durations than previews degraded by 10% (recall that if decreasing degradation of the previews leads to faster reading times, this suggests that the preview facilitated lexical identification). As such, this runs counter to the suggestion that increasing levels of degradation should reduce processing costs (Marx et al., 2015). Within skilled adult readers of English, it would appear that degraded previews do not reduce processing costs.

In sum, the choice of parafoveal mask (letter mask preview vs. degraded preview) does merit attention. For adult readers of English, though, degraded previews do not appear to reduce processing costs in comparison to parafoveal letter masks (Vasilev et al., 2018), suggesting that the use of parafoveal letter masks is still a viable and useful technique.[2] In addition, the use of parafoveal letter masks with developing child readers might lead to overestimation of preview benefit (Marx et al., 2015), but, critically, this has no bearing on

---

[2] That is provided display change awareness is taken into account, as this can affect fixation durations (e.g., Slattery et al., 2011; White et al., 2005). Typically, as is the case in the experiments reported within this thesis, participants who report noticing anything strange about the sentences they read are not included in data analyses.

the effects that can be found when comparing experimental preview manipulations. Consequently, parafoveal letter masks can still provide useful insights into parafoveal pre-processing, and what information readers are able to extract from word *n*+1.

1.2.2.4 Summary

Evidently, like for adult readers, children's eye movement behaviour during reading, regarding *when* they move their eyes, is primarily determined by lexical processing (e.g., word frequency effect; Blythe et al., 2009). Indeed, the speed with which lexical processing can be completed seems to be the main factor behind the differences that are present between adults' and children's eye movements during reading (e.g., Blythe et al., 2009, 2011, 2015; Joseph et al., 2009, 2013; Mancheva et al., 2015; Pagán et al., 2021; Reichle et al., 2013; Tiffin-Richards & Schroeder, 2015b). The reduced rate of lexical processing within child readers could be due to their lexical representations being of lower quality than those of adult readers (e.g., Luke et al., 2015; Perfetti, 2007), causing lexical identification to be more difficult and, therefore, slower. Given the greater difficulties beginner child readers appear to face in their foveal processing, it is unsurprising then that fewer processing resources can be allocated to upcoming information (i.e., their perceptual/letter identity spans are smaller than those of adult readers; e.g., Häikiö et al., 2009; Rayner, 1986). Developmental change is present though, such that, typically, with age and reading skill, more attentional resources can be devoted to information beyond the point of fixation (e.g., Häikiö et al., 2009; Sperlich et al., 2015). As such, parafoveally pre-processing would seem to be a skill that gradually develops, and has been shown, to a limited extent in comparison to the literature on adults, to be facilitative to children's lexical identification. Clearly, the ease with which lexical identification can be, and is, achieved by skilled adult and beginner child readers plays a critical role in determining their eye movements. There are still clear gaps though in our understanding of what information child readers of English are able to process from word *n*+1, facilitating their lexical identification, and how this skill might develop.

1.3 General Summary and Overview of Present Thesis

Overall, both the moving window paradigm (McConkie & Rayner, 1975) and the boundary paradigm (Rayner, 1975) have provided interesting insights into the perceptual/letter identity spans, and the consequent parafoveal pre-processing abilities, of adults and children alike. Similar to children's eye movement behaviour being subject to changes with age, driven by improvements in lexical processing speed (e.g., Reichle et al., 2013), their perceptual/letter identity spans also undergo age-related changes, and are modulated by ease of processing (e.g., reading skill and text difficulty; Häikiö et al., 2009; Rayner, 1986), allowing increasingly more attention to be allocated to the parafovea. More research is needed: 1) to examine how speed of lexical processing might affect the time course with which information can be extracted from word *n*+1; and 2) to further examine what characteristics of word *n*+1 can be extracted by 8- to 9-year-old child readers of English. In short, more research is needed in order to examine how this hallmark ability of fluent, skilled adult reading develops.

It is evident that orthography (as outlined in this chapter) and phonology (as outlined in Chapter 2) play key roles in determining the ease with which both adult and child readers are able to lexically identify word *n*. Moreover, research has shown that both orthographic and phonological information can be extracted from word *n*+1 by skilled adult readers of English, facilitating their lexical identification. Clear questions remain, however, as to what information beginner child readers of English are able to extract from word *n*+1, and how parafoveal pre-processing might undergo developmental change. My research findings, reported within this thesis, advance the current knowledge regarding parafoveal pre-processing in child readers of English, by further examining what characteristics of word *n*+1 are processed: the roles orthography and phonology play in preview, and what roles they might play in driving the first-letter bias. Here, I briefly outline the contents of each chapter, and my contribution to the published (or to be published) manuscripts.

In Chapter 2 (Milledge & Blythe, 2019) the literature regarding phonological processing during reading is reviewed. Orthography and phonology are tightly linked within alphabetic languages: a given word's printed form acts as a gateway, allowing readers to access the word's associated speech sounds. Current theories of reading development propose that a reader's reliance on these two sources of information changes; that is, phonology becomes less important, and there is a shift to the predominant use of orthographic information. This supposition, though, is not supported by eye movement research. Such research suggests that phonological processing continues to play a role throughout reading development, as beginner child readers progress to be skilled adult readers, but the nature of this processing changes. A core aspect of this developmental change in phonological processing is the ability to extract phonological information from word *n*+1.

In order to determine whether beginner child readers of English are able to process phonological information from preview, thereby demonstrating this developmental change, it needed to first be determined whether children can extract orthographic information from the whole of word *n*+1. Subsequently, it could then be determined whether 8- to 9-year-old children have the adult-like ability to use phonology pre-lexically, such that it facilitates lexical identification. It is only on the basis of filling these gaps in knowledge that the interplay between orthography and phonology within children's parafoveal pre-processing could begin to be studied. Moreover, by comparing children's parafoveal pre-processing to that of adults, insight could be gained into differences in the time course of their processing.

Consequently, in the first experimental chapter (Chapter 3; Milledge et al., 2020), a study is reported where the extent to which sensitivity was shown by child readers to whole-word orthography in preview, and the importance of external letters compared to internal letters to lexical identification was examined. It was found that child (and adult) readers were able to extract orthographic information from a whole word in preview, albeit slower within the children. It was also found that external letters in preview were more facilitative to lexical identification than internal letters in preview. In addition, regarding the external letters, the first letter played a particularly important role.

Then, given adult and child readers displayed similar effects with regard to pre-processing of orthography (i.e., sensitivity to whole-word orthography in preview), the developmental change in phonological processing was directly examined in Chapter 4 (Milledge et al., 2021a). The results show that the children, like the adults, were able to benefit from phonological information being provided for word n+1; phonology facilitated their lexical identification. Moreover, orthography was found to affect the extent to which both adult and child readers were able to benefit from phonological information in preview. Again, the results highlight the importance of the first letter to lexical identification.

Finally, on the basis of the results found within Chapters 3 and 4, Chapter 5 (Milledge et al., 2021b) examined the interplay between orthography and phonology with regard to the importance of the first letter to lexical identification (first-letter bias). The results indicate that orthographic encoding of the first letter is crucial to the efficient processing of phonological information from word n+1 for both skilled adult and beginner child readers. Evidence, again, was found that children's extraction of orthographic information from preview was slower compared to that of the adults.

The present research extends the literature on parafoveal pre-processing in two ways. First, by examining novel areas of parafoveal pre-processing within beginner child readers of English. Second, by examining, and comparing such processing to, skilled adult readers of English; thereby providing some insight into potential developmental change within parafoveal pre-processing.

The preparation of all manuscripts was performed primarily by me, with revisions made on the basis of supervisory feedback and reviewers (in the case of published manuscripts). Experimental ideas were formulated in collaboration with my supervisors. Both the novel experimental stimuli within Chapter 3 and the stimuli within Chapter 5 were primarily of my design. All data collection (including pre-screening of the stimuli used within Chapters 3 and 5) and analyses were conducted by me; though, of course, guidance was provided by my supervisors throughout.

Chapter 2 The Changing Role of Phonology in Reading Development

2.1 Abstract

Processing of both a word's orthography (its printed form) and phonology (its associated speech sounds) are critical to lexical identification during reading, both in beginning and skilled readers. Theories of learning to read typically posit a developmental change, from early readers' reliance on phonology to more skilled readers' development of direct orthographic-semantic links. Specifically, in becoming a skilled reader, the extent to which an individual processes phonology during lexical identification is thought to decrease. Recent data from eye movement research suggests, however, that the developmental change in phonological processing is somewhat more nuanced than this. Such studies show that phonology influences lexical identification in beginning and skilled readers, in both typically and atypically developing populations. These data indicate, therefore, that the developmental change might better be characterised as a transition from overt decoding to abstract, covert recoding. We do not stop processing phonology as we become more skilled at reading; rather, the nature of that processing changes.

2.2 The Changing Role of Phonology in Reading Development

Learning to read is a vital process within modern societies, given how much information is conveyed by the written word, ultimately affecting academic success, employability, and social and economic welfare. For example, it is estimated that the cost of illiteracy to the global economy is over $1 trillion each year; costing a developed nation 2% of its gross domestic product (GDP), an emerging economy 1.2% of its GDP, and a developing country 0.5% of its GDP (World Literacy Foundation, 2015). Yet the acquisition of this skill, so pivotal to successful functioning within society, is a long, complicated, and effortful process that can last for many years.

Reading is a process that requires the learning of associations between the visual forms of printed words (orthography) and their associated speech sounds (phonology) and meanings (semantics). The aim of reading is to construct meaning from text, i.e., for the reader to comprehend the written language. It is well-recognised, though, that making these links from orthography to semantics also involves phonological processing (Leinenger, 2014). Oral language acquisition precedes written language acquisition, and so a child's earliest cognitive representations of words include phonology and semantics; only later, as they learn to read, do those phonological and semantic representations map onto orthographic forms (Frost, 1998).

Within theoretical accounts of reading development, a broad consensus seems to be that as a child's reading skill increases, their lexical identification becomes increasingly based on direct orthographic-semantic links and the contribution of phonology to lexical identification decreases (e.g., Castles et al., 2018; Ehri, 1995, 1998, 1999, 2005, 2007; Frith, 1985). Consequently, skilled reading is often characterised as an individual's ability to access semantics directly from a word's printed form. This view has been supported by data from pen-and-paper tasks, such as hand-coding of a child's reading, spelling, or pronunciation errors (e.g., Adams & Huggins, 1985; Doctor & Coltheart, 1980; Greenberg et al., 1997). In recent years, though, eye movement research has indicated that children continue to process phonology during lexical identification as their reading skills increase (e.g., Blythe et al., 2015). These data indicate that developmental change in phonological processing is better characterised as a progression from early, overt decoding (the conscious, effortful sounding out of printed letters to identify a word) to more sophisticated, covert phonological recoding (the rapid, covert, pre-lexical processing of a printed word's phonology).

We begin by briefly reviewing the literature on theoretical models of children's reading development, which clearly documents a developmental change in phonological processing during lexical identification. We then review the literature on skilled adult readers' lexical identification which has examined, in considerable detail, the role of phonological processing. Subsequently, research within developmental populations, both typical and atypical, is discussed. Phonological processing in languages other than English is also briefly considered.[3] Finally, some models of word recognition will be briefly outlined and then evaluated within the context of this paper. Taken together, we consider how these recent contributions to the experimental literature might contribute to both theoretical models of learning to read and models of word recognition.

2.3 Theories of Learning to Read

One prominent theory of how visual word recognition skills develop is Share's (1995) self-teaching hypothesis. This hypothesis posits that phonology plays a central role in how readers acquire orthographic representations of words. Phonological decoding (to achieve a correct pronunciation) is assumed to be critical for the acquisition of orthographic representations, as it draws the child's attention to the order and identity of a word's

---

[3] Given how theories of learning to read relate primarily to reading development within English, this paper's focus will predominantly be on research conducted in English.

constituent letters. As such, decoding provides children with the opportunity to set up direct connections between the spelling of a letter string and the phonology of the spoken word, which results in the growth and development of their lexicons. In this way, phonology serves as a powerful self-teaching device: the explicit learning of a few sets of grapheme-to-phoneme correspondences (GPCs) allows children to decode an increasing number of words, which, in turn, supports the growth of their lexicons.

A number of theories have been proposed in order to try to characterise the process that children go through as they progress from beginning to skilled reader, with many proposing that children progress through a series of phases (e.g., Ehri, 1995, 1998, 1999, 2005, 2007; Frith, 1985; Gough & Hillinger, 1980; Stuart & Coltheart, 1988) as they become more experienced in dealing with written text, ultimately leading to fluent, skilled reading. It is assumed that whilst most children pass through these phases, they are not biologically determined (Rayner et al., 2012). These phases are described as representing the reader's dominant (but not sole) process for identifying words during reading, at that point in the child's development. There are, of course, differences between the theories of reading development. For example, some theories suggest three phases (e.g., Frith, 1985) whilst others suggest four phases (e.g., Ehri, 1995, 1998, 1999, 2005, 2007). Here we focus upon the common aspects that are relevant to our interest in phonological processing. Broadly speaking, the earliest phase(s) of reading development is characterised by a child's attempts to learn associations between orthographic features of written text (although not complete word forms) and words that already exist in their oral vocabulary (e.g., recognising the word *camel* because it has two humps in the middle; Gough et al., 1992). Subsequently, children learn the alphabet and, consequently, learn grapheme-to-phoneme correspondences (e.g., learning that the word *cat* is pronounced */k/ /æ/ /t/*), providing the capability to read words the child has not encountered before. Then, finally, a child progresses to the point where they are able to identify the majority of printed words that they encounter through whole word recognition, with the assumption that this process relies on direct orthographic-semantic links. At this point, a child does not engage in any observable, overt phonological decoding in order to identify words during reading (for a recent review see Castles et al., 2018).

A major similarity between these theories of reading development is that they propose a developmental shift from beginning readers relying more on phonology to identify words, to more skilled readers forming direct links between orthography and semantics (e.g., Ehri, 1995, 1998, 1999, 2005, 2007; Frith, 1985). Inherent in this proposed trajectory is the decreased reliance on phonology, to the point where it no longer contributes to lexical identification for most words that a reader encounters. Such theories, though, were primarily formulated on the basis of findings from off-line tasks (e.g., Doctor & Coltheart, 1980; Ehri & Wilce, 1983; Rosenthal & Ehri, 2008). Whilst it is true that off-line tasks, and isolated word recognition tasks (as discussed in the following section), have provided researchers with insight (albeit indirect) into the role phonological processing plays in both skilled adult and beginning child readers, and the shift from effortful phonological decoding to fluent sight word reading (e.g., Ehri, 2005), it is eye movement research (discussed in Section 2.5) with skilled adult readers, and more recently with developmental populations, that has provided direct insight into how this proposed theoretical developmental shift may be more nuanced than these current theories account for.

2.4 The Role of Phonology: Isolated Word Recognition Tasks

This section will outline four key areas of evidence: delineating how isolated word recognition tasks have demonstrated the use of overt phonological decoding by beginner readers in order to achieve lexical access; how this subsequently decreases based on reading skill; and how adults display covert phonological recoding; with children also having been found to display this form of phonological processing.

A substantial body of evidence has documented how readers engage in overt phonological decoding in order to identify printed words, using a variety of experimental paradigms. For example, lexical decision tasks (LDTs), where participants are required to decide, as quickly as possible, whether a printed letter string is a real word or not; semantic categorisation tasks, that require the participant to decide whether or not each presented word is an exemplar of a particular semantic category; and naming tasks, that require participants to pronounce a written letter string, often at speed, have all been used.

First, such methods have documented overt phonological decoding in beginning readers. For example, Johnston and Thompson (1989) found that 8-year-old English children were less accurate at rejecting pseudohomophones (e.g., *wotch* - *watch*) than ordinary nonwords (e.g., *cotch*) in a LDT (Experiment 1). It was noted that many of the children tended to sound the stimuli out loud prior to making the lexical decision. Sounding out is a clear indication of phonological decoding being undertaken by the children, and the children displayed reduced accuracy in rejecting the nonword pseudohomophones, indicating that lexical entries were being activated for their respective ‘real word’ homophones. Phonological decoding was enabling the children to activate an existing lexical entry due to shared phonology, regardless of the status of the pseudohomophone as a nonword (with no possible lexical entry). This tendency for children to rely on phonological decoding seems to become particularly apparent when they encounter unfamiliar words. For example, Adams and Huggins (1985) selected 50 exception words, such as *ocean*, *sword*, and *yacht*, which were ordered by frequency (how often a word is typically encountered in text), so that easier words preceded harder words. The researchers found that children in Grades 2-5 typically read words accurately and without any overt decoding until they reached a point in the list where the words became unfamiliar (i.e., low frequency words). At this point readers began sounding out and blending the words, which caused them to hesitate and often misread the words. Schmalz et al. (2013) found that children showed regularity effects (whereby a benefit is found for regular words, that is words with pronunciations that conform to GPC rules, e.g., *spade*, over irregular words, with pronunciations that do not conform to GPC rules, e.g., *yacht*) for low frequency words (e.g., *desk* vs. *calm*) but not high frequency words (e.g., *mess* vs. *ghost*) on a LDT. The researchers argued that children were using phonological decoding for words that they encounter less frequently because the output for irregular words from phonological decoding conflicts with the correct entry in the mental lexicon. For high frequency words, however, the lack of regularity effects suggests that children as young as 8-years-old were relying predominantly on a direct route from orthography to semantics for high frequency words.

Second, the literature shows children's decreasing reliance on overt phonological decoding as their reading skill increases. It is posited that readers increasingly identify words by sight, with direct links from orthography to semantics (e.g., Ehri, 2005). For example, Samuels et al. (1978) used a semantic categorisation task with children from Grades 2, 4, and 6, as well as college students. The words used in this task varied in length from three to six letters. Whilst second graders’ response latencies increased as words grew longer, older students’ latencies did not change as a function of word length. This suggests that the older participants were processing the words as wholes, whilst the second graders were processing component letters in order to read the words (although it is worth noting that this could be an orthographic effect rather than an effect of phonology). Nevertheless, other research has also demonstrated how phonological decoding decreases as reading skill increases. For example, Ehri and Wilce (1983) measured the latencies of skilled and less skilled readers (from Grades 1-4) in a series of naming tasks using common words (e.g., *book*), number words (e.g., *four*), CVC nonwords (e.g., *jad*), and single digits (e.g., *6*). Skilled readers, across the Grades, named words faster than nonwords, and named words as quickly as digits, indicating that they were processing the words as wholes. In contrast, though, the less skilled readers only displayed this pattern of effects in Grade 4; only the oldest less skilled readers were equally as fast naming words as digits. Overall, these data show that as children become increasingly skilled readers, decoding decreases. Researchers have often inferred from this an increasingly dominant process of direct access from orthography to semantics.

Third, a large body of evidence has documented phonological recoding in skilled adult readers. For example, Lesch and Pollatsek (1993) had participants name target words (e.g., *nut*) after the presentation of a prime, either a semantic associate word (e.g., *beech*), a homophone of that associate (e.g., *beach*), or an orthographic control (e.g., *bench*). The researchers found that, at short prime durations, the target words were named faster following both the semantic associates and the homophone primes, in comparison to the orthographic controls. The researchers concluded, therefore, that phonological recoding contributed to readers' lexical access. Van Orden (1987), in a semantic categorisation task, found that frequent errors were made to homophones of particular categories; for example, for the category 'flower' the word *rows* is homophonic to the category instance of *rose*, and participants frequently made false positive errors to *rows*, relative to orthographic controls (e.g., *robs*). As such, phonology appears to play an important role in allowing adults to achieve lexical access through phonological recoding (e.g., Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994a, 1994b; Rubenstein et al., 1971; Van Orden, 1987).

Fourth, children have also been shown to display phonological recoding, with this form of processing seeming to be pivotal in the development of visual word recognition skills. For example, Kyte and Johnson (2006) had Grade 4 and 5 children make lexical decisions for monosyllabic words (e.g., *bean*/*meat*) and pseudowords (e.g., *meap*/*meep*) under two matched experimental conditions: one where items were named prior to lexical decision to promote phonological recoding (read aloud condition), and a condition presumed to limit phonological recoding (concurrent articulation condition; participants repeated a syllable (e.g., 'LA') whilst completing the LDTs). Later, approximately 24 hours after the LDTs, orthographic learning of the pseudowords was evaluated using orthographic choice, spelling, and naming tasks. Target words learned with phonological recoding produced greater orthographic learning than those learned with concurrent articulation. This study provides some evidence for the importance of phonological processing in the development of visual word recognition skills and an orthographic lexicon (consistent with the self-teaching hypothesis; Share, 1995). However, it is important to note that this task requires overt phonological processing in order to name each stimulus aloud; such processing is not required in silent sentence reading. Error detection tasks have also been used to examine phonological recoding in children, where participants are required to decide whether an error is present in the context of a whole sentence. For example, Coltheart et al. (1988) asked adults (Experiment 1) and children (Experiment 2) to judge whether printed sentences were correct or not. One of the unacceptable sentence conditions presented pseudohomophones (e.g., *Her bloo dress was new.*). The researchers argued that, in this condition, any observed effects of phonology must be pre-lexical because there are no lexical entries for nonwords (i.e., it is not possible for phonology to have a top-down influence, post-lexical access, as could be the case for known words). Pseudohomophone sentences resulted in significantly higher false positive rates for both adult and child readers, relative to control conditions. Thus, the authors argued that both the adults and the children were pre-lexically processing phonology (recoding). One possible caveat is that response times were not recorded, only accuracy. It is possible that readers were engaging in some form of subvocal phonological decoding in order to process the pseudohomophones.

Taken together, these studies provide strong evidence for phonological recoding in skilled adult readers (e.g., Lesch & Pollatsek, 1993; Lukatela & Turvey, 1994a, 1994b; Van Orden, 1987). There is also clear evidence that beginning readers rely on phonological decoding, and that this reduces over time as reading skill increases (e.g., Ehri & Wilce, 1983; Samuels et al., 1978). Finally, there is some evidence that once children are past the point in their reading development where they are engaging in effortful phonological decoding, they have made a transition to phonological recoding (e.g., Coltheart et al., 1988; Kyte & Johnson, 2006). Whilst these studies do suggest such a transition, they do not afford as direct insight into reader's cognitive processing of text as eye movement research does, especially given the off-line nature of some of the data (e.g., Coltheart et al., 1988). Consequently, seeking converging evidence from different approaches could prove useful.

2.5 The Role of Phonology: Eye Movement Research

Eye movement research provides a highly sensitive index of cognitive processing during reading, affording researchers an insight into the on-line, moment-to-moment, operations involved in the reading process (Rayner, 1998, 2009; Starr & Rayner, 2001). As such, researchers can gain insight into the cognitive processing of text using more naturalistic sentence reading, as opposed to isolated word recognition tasks or off-line tasks. A body of literature has used eye movement recordings to examine the contribution of phonological processing to lexical identification during silent sentence reading.

2.5.1 Adults

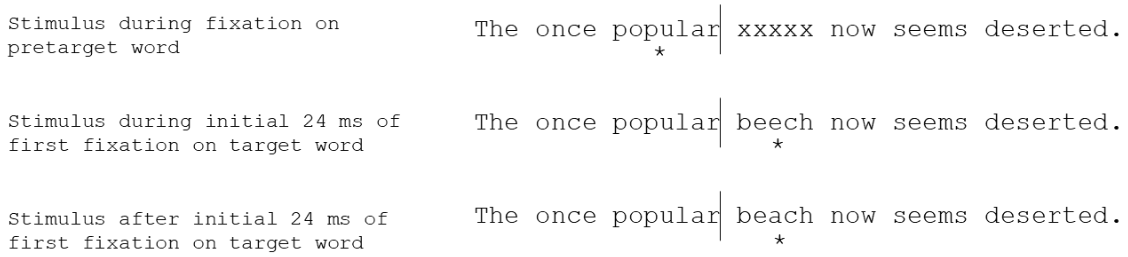
Research has strongly indicated that adults continue to make use of phonology during reading. From the literature on skilled adult reading, two roles have been proposed for phonology during skilled reading: (1) phonology may play a pre-lexical role, and aid the process of lexical access and word identification; or (2) phonological codes may be activated as a function of lexical access or after lexical access (Frost, 1998; Leinenger, 2014).

Rayner et al. (1998) provided evidence that phonological information is activated during silent reading. Participants read short passages that contained a correct target word, a homophone, or an orthographic control (e.g., *Murderers who kill many people according to a pattern are referred to as serial/cereal/verbal killers.*). Both the orthographic controls and the homophones were incongruent with the semantics of the sentence context and, as such, longer reading times would be expected in both these conditions, relative to the correct target word. Importantly, the orthographic controls and homophones were matched in terms of their orthographic overlap with the target word. Shorter reading times on the homophone relative to the orthographic control would, therefore, be attributable to the homophone's shared phonology with the correct target word. Strikingly, reading times on the homophone were not significantly different from reading times on the correct target word, when it was orthographically similar to the target word (e.g., *heal*-*heel* vs. *right*-*write*). This suggests that readers' early activation of congruent phonological codes resulted in the reader not even noticing that the word they were fixating was an error word (that is, a word that was incorrect in the context of the sentence). Critically, across both orthographically similar and dissimilar conditions, participants displayed shorter reading times on homophones than on orthographic controls, and this effect was observed in early measures of processing (i.e., in first fixation duration- the duration of the first fixation on a word regardless of how many fixations it receives). It is worth noting that in the researchers' first experiment a pseudohomophone condition (e.g., *brane* - *brain*) was also used, and the pattern of results was similar to that of the homophones. This provides further evidence for a pre-lexical role for phonology: pseudowords do not have lexical entries, so any characteristics of such words that facilitate lexical identification (i.e., shared phonology with real words) would have to be activated before lexical access is achieved (Lukatela & Turvey, 1994a). Thus, phonological recoding was used by skilled adult readers in their initial fixation on a word, seemingly pre-lexically, facilitating lexical identification.

With respect to the pre- versus lexical/post-lexical phonology question, though, the strongest evidence comes from fast priming (Figure 2.1; Sereno & Rayner, 1992) and parafoveal pre-processing studies.

Figure 2.1

An example of the fast priming technique. The asterisk underneath each sentence indicates the reader's fixation location. An invisible boundary is placed in a sentence, in the space before a target word (the lines in the example below represent the location of the boundary, but this is not visible to participants). Before fixation a string of 'x's is present where the target word should be. When the readers' eyes cross the invisible boundary, and first fixate the target word location, a prime is presented for a very brief amount of time (e.g., 24 ms), before being replaced by the target word. This example shows a homophone prime (e.g., beech) for a target word (e.g., beach).



Stimulus during fixation on pretarget word — The once popular xxxxx now seems deserted.

Stimulus during initial 24 ms of first fixation on target word — The once popular beech now seems deserted.

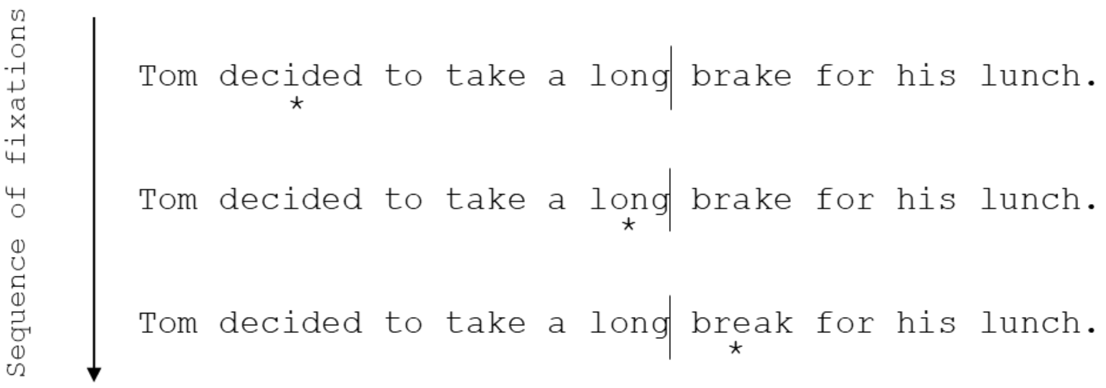
Stimulus after initial 24 ms of first fixation on target word — The once popular beach now seems deserted.

Rayner et al. (1995) used the fast priming technique to compare identity (e.g., *beach*), homophone (e.g., *beech*), orthographic control (e.g., *bench*), or dissimilar primes (e.g., *noise*). The critical comparison here was that of reading times on the target word when it was primed by a homophone relative to an orthographic control (i.e., looking for evidence of a phonological priming effect). Participants had shorter gaze durations on a target word when it had been preceded by a homophone prime than when it had been preceded by an orthographic control. It would appear, then, that phonology can be coded quickly enough to facilitate lexical access and identification of the target word. Further evidence for this argument is provided by parafoveal pre-processing studies.

Parafoveal pre-processing refers to readers' extraction of information from the next word in a sentence (referred to as *n*+1) before it is directly fixated (whilst processing is ongoing for the currently fixated word- referred to as *n*). It is typically investigated using the boundary paradigm (Figure 2.2; Rayner, 1975).

Figure 2.2

An example of the boundary paradigm. The asterisk underneath each sentence indicates the reader's fixation location. An invisible boundary is placed in a sentence, in the space immediately before a target word (the lines in the example below represent the location of the boundary, but this is not visible to participants). A preview letter string is available in the target word's location prior to the reader making a saccade that crosses this invisible boundary. After the reader's eyes cross the boundary and they move to directly fixate the target word, then a display change occurs wherein the preview letter string changes to the correct target word. By manipulating certain characteristics of the overlap (e.g., phonological similarity) between the preview string and the target word, parafoveal pre-processing can be studied. For example, phonologically consistent (e.g., brake) and inconsistent (e.g., bread) previews can be presented for a target (e.g., break) to examine the extent to which a reader is pre-processing phonology prior to direct fixation. If a reader does extract phonological information during parafoveal pre-processing, then reading times on the target word should be shorter following a consistent preview than an inconsistent preview. This decrease to reading times is referred to as preview benefit. If preview benefit is found, i.e., shorter reading times, on a word that was parafoveally available compared to when the parafoveal preview word was masked, this is strongly indicative of parafoveal pre-processing having occurred, as lexical identification has been facilitated. As such, parafoveal pre-processing, and this paradigm, enables researchers to investigate pre-lexical effects, as manipulations are conducted outside of direct fixation (i.e., lexical processing): if the manipulated characteristic of a given word in the parafovea confers preview benefit *to the reader, the word must have been pre-lexically processed to some extent prior to it receiving a direct fixation.*



Indeed, evidence from the use of the boundary paradigm has found that phonological recoding begins prior to direct fixation in skilled adult readers. For example, Pollatsek et al. (1992) found that readers can pre-process phonological cues from an upcoming word. Previews were either homophones or orthographic controls for a target word that was presented after the reader's eyes had crossed the boundary. They found that reading times on the correct target word were shorter when the preview had been a homophone than when it had been an orthographic control. Such effects, indicating pre-lexical parafoveal processing of phonology, have now been shown in a number of studies looking at parafoveal pre-processing in skilled adult readers (e.g., Ashby et al., 2006; Henderson et al., 1995; Jouravlev & Jared, 2018; Pollatsek et al., 1992) and the fast priming technique has provided similar findings (Rayner et al., 1995). This suggests that phonological recoding plays a key role in activating lexical entries during skilled adult reading; that is, a word’s phonology plays a pre-lexical role rather than a lexical/post-lexical role.

2.5.2 Children

Far less research has been done with children using research methods that are sensitive to on-line cognitive processing during reading. To date, though, two studies have used eye movements to examine phonological processing during children's silent sentence reading, examining foveal reading processes. Blythe et al. (2015) presented sentences containing correct target words, pseudohomophones, or orthographic controls, to both adults and children aged 7- to 9-years (e.g., *Today we had a huge water/worta/wecho fight in my back garden.*).[4] They found that children, similar to adults, benefitted from the valid phonology of a pseudohomophone compared to an orthographic control. These data were argued to provide evidence for covert phonological recoding in children as young as 7-years-old (contradictory to some isolated word recognition research; e.g., Johnston & Thompson, 1989). Two further points support this conclusion. First, all participants were reading silently, and no overt decoding was observed at any point. Clearly, these children were beyond the phase of reading development where overt decoding was their primary strategy for lexical identification. Second, and critically, when compared against reading times on the correct target word within a sentence, the cost associated with pseudohomophones was less than 200 ms, and reading times on the pseudohomophones was less than 600 ms in total in the children's data. These reading times are too short to plausibly incorporate the sounding out and then blending together of phonemes. These data are, therefore, most consistent with phonological recoding during lexical identification, suggesting that both adults and children were able to access the correct lexical representation on the basis of a letter string's phonology.

Moreover, Jared et al. (Experiment 3; 2016) provide further evidence that phonological representations are used in the initial activation of word meanings. The researchers monitored children's (10- to 11-year-olds) eye movements as they read sentences silently,

---

[4] Pseudohomophones were used due to this age group of children being limited in the number of homophone pairs known to them, especially with Age-of-Acquisition limited to earlier than 6 years to maximize the likelihood that all participants would be familiar with the target words.

some of which contained a correct target word (e.g., *whether*), some a homophone (e.g., *weather*), and some an orthographic control (e.g., *winter*). Critically, the homophones were not as disruptive to the children's reading as the orthographic controls (i.e., the children displayed shorter reading times when a homophone was present than when an orthographic control was present). This observed homophone advantage reflects the contribution phonology made to activating the meanings of words for the child readers (regardless of word frequency). Phonology, therefore, seems to play a key role during children's lexical identification during silent sentence reading. Furthermore, similar to Blythe et al. (2015), the mean reading times suggest that children were undertaking phonological recoding (as opposed to overt decoding).

This research (Blythe et al., 2015; Jared et al., 2016) is consistent with the view that phonology continues to play a role in aiding lexical access, but in an increasingly covert manner as age and reading skill increase (Castles et al., 2018; Ehri & Wilce, 1985). This argument is further supported by studies that have shown increased fixation times on long words (e.g., *medicine*) compared to short words (e.g., *salt*) in both children and adults (Hyönä & Olson, 1995; Joseph et al., 2009). There are two critical points to note with respect to these studies. First, Hyönä and Olson (1995) used a reading aloud task with 8- to 12-year-old children, and no overt decoding was observed for either the long or the short words. Second, the magnitude of the increase in reading times was between 22 ms per letter (Joseph et al., 2009; silent reading in 7- to 11-year-old children) and 58 ms per letter (Hyönä & Olson, 1995). The magnitude of these increases to reading times are too small to conceivably argue that children were sounding out and blending phonemes together, either vocally or subvocally, in order to achieve lexical access (phonological decoding). Both of these points support the argument that children at this age have moved beyond overt phonological decoding during lexical identification.

It is widely recognised that adults continue to make use of phonology to aid lexical access and identification during reading (e.g., Frost, 1998; Leinenger, 2014), but until recently this issue has been somewhat neglected within the empirical literature on children's reading development. We contend that, while there is developmental change in phonological processing during reading, this is best characterised as a transition from phonological decoding to phonological recoding. Such a developmental transition is not currently accounted for in theoretical models of learning to read, which simply posit decreasing reliance on phonology as reading skill increases (e.g., Ehri, 1995, 1998, 1999, 2005, 2007; Frith, 1985).

It is worth noting that phonological processing in English, the focus of this paper, may differ from that in other languages, due to differences in orthographic depth (the consistency of a language’s GPCs). For example, English has an opaque orthography, wherein GPCs are not very consistent (i.e., *ough* in *cough*, *through*, *though*, etc.), whilst other alphabetic languages, like Finnish and German, benefit from more transparent orthographies. One piece of research has investigated phonological pre-processing in German. Tiffin-Richards and Schroeder (2015a) found that German adults benefitted more from orthographic, than phonological, information in the parafovea. Whilst children also gained some preview benefit from orthographic information in the parafovea, this was only under certain conditions: when the target words only received a single fixation and when capitalisation of the word was present. In contrast, the children did show clear processing benefit from pseudohomophones. This would suggest that, in German, for children phonology plays a more important role in word identification than orthography, whilst, for adults, the opposite pattern seems to occur: orthography seems to play a more dominant role in facilitating lexical access than phonology. In Chinese, a morphosyllabic language (Perfetti et al., 2005), phonological information has been shown to be activated pre-lexically by children, whilst adults seemed to use more direct access from orthography to semantics (Zhou et al., 2018).

Within Chinese, the researchers argued, early, pre-lexical activation of phonology diminishes as readers become more skilled. It is worth noting though that this research focuses on parafoveal processing of orthographic and phonological information and, so, does not make claims that, for instance, children do not process orthographic information foveally in German. Overall, though, this research on both adults' and children's parafoveal pre-processing in German and Chinese seems to be in contrast to the research looking at pre-processing of phonology in English adults (e.g., Pollatsek et al., 1992). Indeed, concerns have been raised as to whether research conducted in English may have overestimated the importance of phonology (e.g., Share, 2008; Ziegler et al., 2010). Consequently, the developmental transition from overt, effortful phonological decoding to covert, rapid phonological recoding that appears to occur in English, as outlined in this paper, may not be applicable to other languages. Whilst phonology does seem to play a role in reading development in other alphabetic languages besides English, it does seem to be modulated by orthographic depth (Ziegler et al., 2010). Evidence suggests that readers of more transparent orthographies might make the transition from phonological decoding to phonological recoding at a faster rate than readers of English, with it suggested that the difficulty associated with progressing to phonological recoding is specific to English and its complex GPCs (e.g., Aro & Wimmer, 2003). Thus, the extent to which reading development within different languages is determined by phonological processing may differ.

2.5.3 Atypical Development

Most recently, studies have begun to show evidence for pre-lexical phonological processing in populations with atypical reading development; specifically, individuals with permanent childhood hearing loss (PCHL; Blythe et al., 2018), and individuals with developmental dyslexia (Blythe et al., 2020). Both of these participant populations are known to commonly experience substantial difficulties in learning to read, and one component of these difficulties is thought to be poor phonological processing skills (e.g., Mayberry et al., 2011; Snowling, 1981; Vellutino et al., 2004).

In the case of individuals with PCHL, their auditory perception since birth has been substantially impoverished and it is likely that this results in underspecified cognitive representations of phonology. Indeed, on tasks that require overt awareness of or conscious manipulation of speech sounds, Blythe et al. (2018) found that teenagers with PCHL scored significantly lower than both chronological age-matched and reading-matched, hearing peers. Despite their difficulties in overt phonological decoding and phonological awareness, these teenagers displayed a pseudohomophone advantage both during direct fixation and from parafoveal preview. In particular, the pseudohomophone advantage shown by teenagers with PCHL was very similar in both time course and magnitude to the effect observed in their younger, reading-matched hearing peers. This strongly indicates that, despite their overall difficulties in learning to read, one particular aspect of lexical processing was maturing in a typical manner (albeit with a slight developmental delay) - the transition to phonological recoding.

In the case of developmental dyslexia, both overall reading difficulties and specific difficulties in phonological awareness and processing have been well-documented; indeed, poor phonological processing skills are largely accepted as the predominant cause of developmental dyslexia (e.g., Snowling, 1981; Vellutino et al., 2004). Again, though, recent research has shown that teenagers with dyslexia still exhibit a pseudohomophone advantage during reading during both direct fixation and parafoveal preview (Blythe et al., 2020). Similar to the data from teenagers with PCHL, this pseudohomophone advantage during silent sentence reading was observed in contrast to significantly poorer performance on overt tasks of phonological processing compared to their typically developing peers.

In sum, eye movement research in recent years has provided strong evidence for pre-lexical phonological recoding by adults, typically developing children, and even individuals with PCHL or dyslexia, during silent sentence reading (e.g., Ashby et al., 2006; Blythe et al., 2015, 2018, 2020; Henderson et al., 1995; Jared et al., 2016; Pollatsek et al., 1992; Rayner et al., 1995). These data challenge theoretical accounts of reading development which posit that phonological processing during lexical identification reduces with time and reading skill (e.g., Ehri, 1995, 1998, 1999, 2005, 2007; Frith, 1985). Rather, these data are more consistent with the view that, as reading skill increases, there is a transition from phonological decoding to phonological recoding. This transition seems to occur relatively early, and to be remarkably robust across both typical and atypical reading development.

2.6 The Role of Phonology: Models of Word Recognition

A number of different models have been put forward by researchers to try to explain how printed word recognition occurs (e.g., the dual-route cascaded model- DRC; Coltheart et al., 2001; the multiple-route model; Grainger et al., 2012; connectionist dual-process model- CDP+; Perry et al., 2007). It is non-controversial that all of these models posit some role for phonology in visual word recognition, but they vary in terms of the importance that is ascribed to phonology. For a recent and comprehensive review see Jared et al. (2016). Here, we briefly outline these models and how each of them incorporates phonological processing into printed word identification. Critically, we consider the degree to which these models can account for developmental change in this respect.

2.6.1 The DRC Model

According to this model, processing is accomplished via two distinct, but interactive, routes: lexical and non-lexical (see Coltheart et al., 2001, Figure 7, p. 214). The lexical (direct) route relies on the activation of word-specific orthographic representations: the features of a word's letters activate the word's letter units (in parallel), and these letters then activate the word's entry in the orthographic lexicon. The non-lexical (indirect) route is based on the use of GPCs (operating serially from left to right); visual features and letter units are activated just as with the lexical route (as they are common to both routes). Processing along the direct lexical route gets faster each time a word is encountered, so the lexical representations of more common words are activated by the direct route before the slower, indirect, non-lexical route has finished processing the word (e.g., Adams & Huggins, 1985; Schmalz et al., 2013). When tested, the DRC was 99% accurate in generating a pronunciation for the 7,981 words in its orthographic lexicon. It can account for many phenomena that are observed in skilled adult reading, including frequency effects, regularity effects, the pseudohomophone advantage, and orthographic neighbourhood effects. With respect to developmental change, however, the model has no learning mechanism, and "…has nothing to say about the actual process of learning to read" (p. 246). The authors argue that it does work well to characterise what a typically developing child reader has learned so far at any point during the process of learning to read, and that young readers (7-year-olds) have reading systems similar to adults, albeit scaled-down versions. It is not clear, however, how the two routes would develop in a beginning reader, or how the model would account for a developmental transition from decoding to recoding.

2.6.2 The Multiple-Route Model

The multiple-route model (see Grainger et al., 2012, Figure 2, p. 282) makes a distinction between the effortful phonological coding of beginning readers and the faster, more automatic, use of phonology that develops with a reader's exposure to print.[5] The initial, overt coding process enables the development of parallel letter processing involving an array of letter detectors that are location-specific (i.e., that encode the locations of letters within a printed word). Two orthographic codes are generated from this: a coarse-

---

[5] Note that what Grainger et al. (2012) refer to as "phonological recoding" is referred to within this paper as phonological decoding.

grained route that allows direct access to semantics; and a fine-grained route that codes the precise ordering of letters within a string, and then activates the corresponding phonemes as well as whole-word orthography. The model clearly predicts strong, phonologically-based effects (e.g., pseudohomophone effects) in younger children, that reduce but do not disappear with increasing age, as the reader transitions to phonological recoding. This model, therefore, seems to be entirely consistent with the experimental observations from the body of published literature reviewed within this paper.

2.6.3 The CDP+ Model

The CDP+ model (Perry et al., 2007), similar to the DRC model, has two processes: a non-lexical one (sublexical route) that links orthography to phonology, that learns GPCs very quickly; and a direct lexical one that links orthography to phonology- orthographic entries are linked to their phonological counterparts (an implementation of the DRC's lexical route). With respect to developmental change, Ziegler et al. (2014) provided a computational simulation of the self-teaching hypothesis (Share, 1995) within the framework of the CDP+ model. They examined the extent to which the model could learn to identify unknown words, based on initial, explicit teaching of key GPC rules and its existing phonological lexicon, similar to what a child might experience. Ziegler et al. (2014) argue that children receive phonics instruction early in their formal education, but they are not explicitly taught the correct pronunciation of every word that they encounter during reading. Rather, as they come across new printed word forms, they use their knowledge of phonics rules to generate a possible pronunciation and determine whether or not this matches with a word that is already represented in their lexicon (through spoken language exposure). This learning loop is referred to as the phonological decoding self-teaching (PDST) hypothesis, and, indeed, the implementation of the PDST hypothesis worked in the context of a real computational model of learning to read (CDP+). Even starting with a small number of GPCs (as beginner child readers would do), the model was able to acquire word-specific orthographic representations for over 25,000 words and read aloud novel words. On the basis of these rudimentary GPCs (and decoding skills), the model could produce pronunciations for unfamiliar words. Despite the opaque orthography of English, the phonological decoding network was still able to learn up to 80% of the words. Overall, phonological decoding seems to serve as a powerful internal teaching device, as implemented in this model, allowing a basic set of GPCs to open up children's (and the model's) abilities to read novel words and gain orthographic knowledge. It is conceivable within the PDST hypothesis that there is a transition from beginner children's phonological decoding to skilled adult readers' phonological recoding, but this has not yet been operationalised.

In sum, all of these models propose that phonology plays a role in visual word recognition. To date, Grainger et al.'s (2012) multiple-route model provides the clearest implementation that might account for the developmental transition from beginner child readers' effortful phonological decoding to skilled adult readers' unconscious, rapid phonological recoding.

2.7 Conclusions

Whilst it is widely recognised that children rely on phonological decoding in early stages of learning to read, current theories do not fully account for skilled readers' pre-lexical processing of phonology, that is phonological recoding (Ehri, 1995, 1998, 1999, 2005, 2007; Frith, 1985), with only one recent model of word recognition seeming to account for this developmental transition (the multiple-route model; Grainger et al., 2012). Eye movement research has shown pre-lexical processing of phonology in typically developing readers from the age of 7-years through to skilled adult readers, as well as in atypical developmental groups, despite the tasks used not requiring any overt phonological processing (Blythe et al., 2015, 2018, 2020; Jared et al., 2016; Pollatsek et al., 1992; Rayner et al., 1998). Thus, eye movement research provides compelling evidence for phonology having a continued, and pervasive, role in facilitating lexical identification during reading (consistent with the multiple-route model; Grainger et al., 2012). As such, recent empirical findings from on-line research methods such as eye movement recordings need to be incorporated into theories of learning to read, and more consideration given to these findings in developmental models of word recognition. In order to accomplish this, further research is needed to understand the nature and time course of the transition from phonological decoding to recoding, examining moment-to-moment cognitive processing during reading in beginning readers.

Chapter 3 Parafoveal Pre-processing in Children Reading English: The Importance of External Letters

The data that supports the findings of this study, and the code used for the main model analyses, are available from:

https://osf.io/fh7rd/?view_only=6044a208c73b4e488361147d7f799874.

3.1 Abstract

Although previous research has demonstrated that for adults external letters of words are more important than internal letters for lexical processing during reading, no comparable research has been conducted with children. This experiment explored, using the boundary paradigm during silent sentence reading, whether parafoveal pre-processing in English is more affected by the manipulation of external letters or internal letters, and whether this differs between skilled adult and beginner child readers. Six previews were generated: identity (e.g., *monkey*); external letter manipulations where either the beginning three letters of the word were substituted (e.g., *rackey*) or the last three letters of the word were substituted (e.g., *monhig*); internal letter manipulations; e.g., *machey*, *mochiy*); and an unrelated control condition (e.g., *rachig*). Results indicate that both adults and children undertook pre-processing of words in their entirety in the parafovea, and that the manipulation of external letters in preview was more harmful to participants' parafoveal pre-processing than internal letters. The data also suggests developmental change in the time course of pre-processing, with children's pre-processing delayed compared that of the adults. These results not only provide further evidence for the importance of external letters to parafoveal processing and lexical identification for adults, but also demonstrate that such findings can be extended to children.

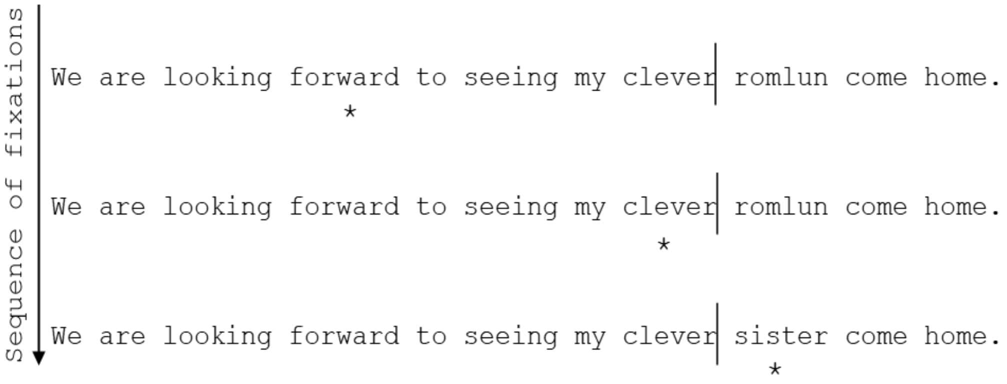
3.2 Introduction

In recent years a number of studies have been reported that examine eye movement behaviour during silent sentence reading in children compared to adults (see Blythe & Joseph, 2011 and Blythe, 2014 for reviews); however, this research has predominantly focused on foveal reading processes. That is, examining word identification processes for the directly fixated word (*n*). In contrast, there is a paucity of research that directly compares parafoveal reading processes in adults and children, examining how identification of the upcoming word (*n*+1) occurs and which factors can affect such processing.

The use of eye movement recordings in order to study reading is a dominant research method for skilled adults; providing a moment-to-moment index of the reader's cognitive processing of text (e.g., Rayner, 2009). Critically, such research has shown that, during a fixation on word *n*, adults both process word *n* and also begin to pre-process word *n*+1. Subsequently, when word *n*+1 is directly fixated, reading times are faster due to the pre-processing that has already occurred (see Schotter et al., 2012 for a review). This is referred to as parafoveal pre-processing, and can be considered a hallmark of skilled, fluent adult reading (Rayner, Liversedge, et al., 2006). The importance of parafoveal pre-processing has been shown through a number of studies that have used gaze-contingent paradigms, where the stimulus changes as the reader progresses through the sentence dependent on the location of their fixation (e.g., the boundary paradigm; Rayner, 1975; see Figure 3.1). Specifically, gaze-contingent techniques can be used to deny readers the opportunity for parafoveal pre-processing. It is quite clear that skilled adult readers depend upon parafoveal pre-processing for rapid, fluent sentence reading.

Figure 3.1

Example of the boundary paradigm (Rayner, 1975). Fixation locations are marked by the asterisk under the sentence. When a sentence is first presented on the screen, the target word is replaced with a preview letter string. When the participant is fixating the pretarget word (n-1; clever in this example), word n (e.g., sister) is unavailable for pre-processing. An invisible boundary is placed immediately in front of the target word (marked here by a vertical line for demonstration, though this is not visible on the participant's screen during the experiment). When the reader makes a saccade across the invisible boundary, the preview letter string (e.g., romlun) is replaced with the correctly spelled word and the reader is typically unaware that any change has occurred. Two control conditions are typically included- an identity condition, where the preview is identical to the target word, and a completely unrelated preview condition, where all letters are replaced with stimulus strings that do not provide any useful information about the upcoming word (e.g., romlun as shown here). Reading times are typically shortest in the identity condition, as the reader has benefitted from undisrupted parafoveal pre-processing of the target word. Conversely, reading times are expected to be longest in the unrelated preview condition, as the reader has been unable to extract any information that might facilitate lexical identification. Experimental conditions then manipulate/preserve features of the upcoming word as per the manipulations of interest in the study. Reduced reading times on a target word observed after a correct (identity) preview, compared to an incorrect preview (i.e., the experimental conditions and the unrelated preview condition), is known as preview benefit.



We are looking forward to seeing my clever romlun come home.

We are looking forward to seeing my clever romlun come home.

We are looking forward to seeing my clever sister come home.

In order to gain insight into how beginner readers progress to be skilled readers, it is crucial to understand how this skill, so pivotal to skilled adult reading, develops. Through the boundary paradigm, by manipulating certain characteristics of the relationship between the preview letter string and the correct target word, it is possible to determine the type of information that is pre-processed in the parafovea. Adults pre-process orthography (a word's printed form), for example displaying faster reading times after an orthographically similar preview is available compared to an orthographically dissimilar preview (e.g., *cahc* vs. *picz* as preview for *cake*; Balota et al., 1985). The external letters of a word are particularly important for skilled adult readers in both parafoveal pre-processing (Johnson et al., 2007) and during subsequent direct fixation (Johnson & Eisler, 2012). Manipulations that affect the first or final letter of a word have a disproportionately large cost to reading times, relative to manipulations of internal letters, with the first letter seeming to play a particularly important role (e.g., Briihl & Inhoff, 1995; Inhoff, 1989a,b; White et al., 2008).

Little research, however, investigating children’s parafoveal pre-processing in alphabetic languages has been undertaken.[6] One study has examined the first letter advantage in parafoveal preview for children compared to adults. Pagán et al. (2016) examined 8- to 9-year-old English children's orthographic pre-processing of the first three letters of an upcoming word. Similar to adults in terms of both the magnitude and the time course of their pre-processing, they also found that children showed a beginning bigram (the first two letters of a word) bias. This study only manipulated the first three letters of words in parafoveal preview, though, and orthographic pre-processing of the entire word form was not examined. Johnson et al. (2018) have also provided evidence for the importance of first letters to children’s pre-processing: faster reading times were found when the first two letters of target words were maintained in preview- orthographically similar condition, compared to when all letters were substituted in preview- orthographically dissimilar condition (e.g., *apydo* vs. *egydo* as previews for *apple*). Thus, the beginning letters clearly play an important role in both adults’ and children’s parafoveal pre-processing, but, whilst Johnson et al.’s (2018) study might suggest that children extract orthography from the entire word in preview, whether children show external letter advantages for both first and final letters or whether this bias is limited to the first letters of a word is unknown.

In the present study two key questions were addressed: (1) whether children are able to pre-process whole target words in the parafovea; and (2) whether external or internal letters are more facilitative to parafoveal pre-processing. To examine these questions, the

---

[6] Two studies have been conducted in English (Johnson et al., 2018; Pagán et al., 2016), one in Finnish (Häikiö et al., 2010), and four in German (Marx et al., 2015, 2016, 2017; Tiffin-Richards & Schroeder, 2015a).

boundary paradigm (Rayner, 1975) was used. The locations of letter substitutions within a word were manipulated to examine the spatial extent of orthographic pre-processing in children compared to adults- letters were substituted in preview at the beginning, middle, or end of the target words. Research using other experimental paradigms has indicated that children do pre-process some information up to 11 character spaces away from the point of fixation, although those studies did not show which lexical characteristics were processed (e.g., word length, word shape, letter identity, etc.; Häikiö et al., 2009; Rayner, 1986; Sperlich et al., 2015). On this basis, we predicted that both adults and children would be sensitive to letter substitutions at the end of the target word as well as at the beginning. We also expected to show a higher cost to both adults' and children's reading from manipulations that involved the first letters of a word, compared to those that involved internal letters within a word (Pagán et al., 2016; White et al., 2008).

3.3 Method

3.3.1 Participants

Forty-two adults (*M* = 22.24 years old) and 42 children (aged 8- to 9-years-old; *M* = 8.76) participated in the eye-tracking experiment. See Table 3.1 for a summary of group characteristics. All had normal or corrected to normal vision, and were native speakers of English with no known reading difficulties. This was confirmed by the reading subtests of the Wechsler Individual Achievement Test II UK (WIAT-II UK; Wechsler, 2005); all participants were within the expected range (adults' composite standardised score range: 99-135; children's composite standardised score range: 104-123; see also Table 3.1).

The number of participants is broadly comparable to other studies on children's parafoveal pre-processing (e.g., Marx et al., 2015; Pagán et al., 2016; Tiffin-Richards & Schroeder, 2015a). The present study is most similar to that of Pagán et al. (2016), but in contrast to their 2 (group) × 7 (condition) design with 56 stimuli, the present study had a 2 (group) × 6 (condition) design with 54 stimuli, as outlined in the following subsection. Consequently, this study had more items per condition per participant (nine items per condition per participant, as opposed to eight).

Table 3.1

Summary of Group Characteristics

		Mean	StDev	*t*	*df*	*p*
Test age (years)	Adults	22.24	3.54			
	Children	8.76	.43	24.49	82	< .001
WIAT word reading	Adults	112.17	4.94			
	Children	111.48	4.40	.68	82	.501
WIAT pseudoword decoding	Adults	111.07	5.84			
	Children	109.67	4.80	1.21	82	.232
WIAT comprehension	Adults	119.52	4.65			
	Children	110.21	5.59	8.30	82	< .001
WIAT composite standardised scores	Adults	122.26	8.13			
	Children	110.95	4.77	7.78	82	< .001

Note. The three right-hand columns give the results of independent samples *t*-tests comparing the adults to the children. The WIAT scores all refer to standardised scores.

3.3.2 Materials and Design

We used the stimuli developed by Pagán et al. (2016) which comprised of 26 target words in sentence frames. These were supplemented by 34 additional target words and sentence frames that we created. Target words were either nouns or adjectives, and were bisyllabic with a CVCCVC structure, with the syllable boundary falling between the second and third consonants (see Table 3.2 for target word properties).

Table 3.2

Linguistic Properties of the Target Words and Sentence Frames

	Target words
Orthographic neighbours (N-Watch; Davis, 2005)	≤ 7
Age of Acquisition (Kuperman et al. 2012)	*M* = 5.81 years *SD* = 1.63
Child frequency counts (Children's Printed Word Database; Masterson et al., 2003)	Range = 3-663 per million *M* = 85 *SD* = 128
Adult frequency counts (English Lexicon Project Database; HAL corpus, Balota et al., 2007)	Range = 0-2,160 per million *M* = 134 *SD* = 324
Understandability (1 *easy* to 7 *difficult*)	Range = 1-1.63 *M* = 1.14
Predictability	Range = .05-.86 *M* = .34

Note. The Ages of Acquisition refer to 50 of the target words, as this information was not available in the database for four of the target words (*conker*, *longer*, *ledges*, and *fences*).

All materials were pre-screened for both the difficulty of the sentences and the predictability of the target words within each sentence, to confirm that the materials were suitable for use with our target age range. For the additional 34 target words, two possible sentence frames were created. Eighty children (8- to 9-year-olds; none of whom took part in the eye-tracking experiment) rated these sentences on a scale of 1 (easy to understand) to 7 (difficult to understand). They also completed a sentence constraint rating (predictability) task for the 94 sentences (as Pagán et al., 2016 did not pre-screen for predictability), where the sentence frame was presented with a blank space in the target location and the children were asked to fill in the word that they thought best completed the sentence. The results from the pre-screening are shown in Table 3.2, and the final stimulus set was selected to ensure that the sentences were easy to understand for our target age range, and that the target word in each sentence was not highly predictable (to minimise skipping). For each of the new target words, one sentence frame was selected for use in the eye movement experiment on the basis of this pre-screening. Six target words and their associated sentence frames were dropped (one from Pagán et al., 2016). The final stimulus set comprised of 54 experimental sentences.

The boundary paradigm (Rayner, 1975) was used. Using this paradigm, the text displayed on the screen changes contingent on where the reader is fixating (see Figure 3.1). A preview letter string occupies the target word location at trial onset but, when the reader makes a saccade to directly fixate the target word (crossing an invisible boundary), the preview letter string changes to the correct target word. In the current experiment, six parafoveal preview conditions (or letter strings) were generated for each target word (see

Appendix A; A.1). There were two control conditions: an identity condition, where the preview was identical to the target word (123456; e.g., *sister* - *sister*), and an unrelated condition, where only the letter shapes of the target word were maintained in preview (dddddd; e.g., *romlun* - *sister*). There were four other experimental conditions which each involved the substitution of three of the letters of the target words in preview: the beginning three letters of each word (ddd456); internal letters 2, 3, and 4 (1ddd56); internal letters 3, 4, and 5 (12ddd6); and the end three letters of each word (123ddd). Both the beginning and end substitution conditions were within one syllable, whilst the middle substitution conditions affected both syllables. Both CVCCVC structure and word shape were maintained in these substitutions. An example stimulus in each preview condition is shown in Figure 3.2.

Figure 3.2

Example of an experimental sentence showing the six parafoveal preview conditions, with the invisible boundary shown (though shown as a visible line here it was not visible to participants). The asterisk refers to a fixation, the target word is shown in bold, and the condition is shown in the brackets.

1. We are looking forward to seeing my clever | **sister** come home. **(123456)**
2. We are looking forward to seeing my clever | **romter** come home. **(ddd456)**
3. We are looking forward to seeing my clever | **somler** come home. **(1ddd56)**
4. We are looking forward to seeing my clever | **simlur** come home. **(12ddd6)**
5. We are looking forward to seeing my clever | **sislun** come home. **(123ddd)**
6. We are looking forward to seeing my clever | **romlun** come home. **(dddddd)**

*

The 54 experimental sentences were counterbalanced across six lists using a Latin Square design (nine sentences per condition per participant). The sentences occupied one line on the screen (maximum = 77 characters; *M* = 60 characters) and each target word was placed near the middle of the sentence.

3.3.3 Apparatus and Procedure

An EyeLink 1000 eye-tracker recorded right eye movements (SR Research). Forehead-and-chin rests were used to minimise head movements. The sentences were presented in 14-point, black Courier New font on the grey background of a 21 in. CRT monitor, with a refresh rate of 120 Hz, at a 60 cm viewing distance; one character subtended .34° of visual angle. Participants were instructed to read normally and for comprehension. Once participants had finished reading a sentence, they pressed a response key, and one-third of the sentences were replaced by a comprehension question, to which the participants responded. After completion of the experiment, participants were asked whether they had noticed anything strange about the appearance of the sentences in the experiment: detecting a display change can affect fixation times (e.g., White et al., 2005). Four adult participants reported noticing something unusual about the sentences, so their data was excluded from the analyses. The whole experiment lasted about 45 minutes per participant.

3.4 Results

All participants scored at least 78% correct on the comprehension questions (adults: *M* = 98%; children: *M* = 92%). The data were trimmed using the clean function in

DataViewer (SR Research).[7] In total 1,886 fixations were merged or deleted (2.36% of the dataset; 693 adult fixations, and 1,193 child fixations).

Reading time data on the target word in each sentence were analysed. Before analysing the local dependent measures, the data were further cleaned: trials in which the boundary change occurred early during a fixation on the pretarget word, and those that occurred late when the display change was not completed until more than 15 ms after onset of fixation on the target word were excluded from the analyses (230 adult trials- 10.14% of the adult trials, and 314 children's trials- 13.84% of the children's trials).[8] Prior to analysis, reading time data were log transformed.

Data were analysed using linear mixed effects (lme) models, using the *lmer* function from the lme4 package (Bates et al., 2015) within the R environment for Statistical Computing (R Core Team, 2020). We focus here upon three dependent

---

[7] Fixations shorter than 80 ms were merged with the neighbouring fixation if within a .50° distance of another fixation over 80 ms, and fixations shorter than 40 ms were merged with neighbouring fixations if within a 1.25° distance of each other. Then if an interest area had three or more fixations shorter than 140 ms, these were merged into longer fixations. Finally, all remaining fixations shorter than 80 ms or longer than 1,200 ms were deleted.

[8] A late boundary change was also operationalised as 10 ms in order to compare the results with the 15 ms report. The pattern of data remained unchanged between the two, so the 15 ms criterion of a late boundary change was used as it allowed the retention of more data (3,992 data points as opposed to 3,837). Regarding the number of items per condition for each participant, after the boundary change cleaning, within the adults the lowest total number of items recorded for a participant was 43 (*M* = 46.52, total range: 42-54; 123456 *M* = 8.00, range: 6-9; ddd456 *M* = 8.02, range: 5-9; 1ddd56 *M* = 8.48, range: 7-9; 12ddd6 *M* = 8.05, range: 5-9; 123ddd *M* = 7.81, range: 4-9; and dddddd *M* = 8.17, range: 6-9) and within the children this was 38 (*M* = 48.52, total range: 38-53; 123456 *M* = 7.79, range: 3-9; ddd456 *M* = 7.43, range: 4-9; 1ddd56 *M* = 7.74, range: 5-9; 12ddd6 *M* = 7.86, range: 5-9; 123ddd *M* = 7.90, range: 6-9; dddddd *M* = 7.81, range: 4-9).

measures: first fixation duration (the duration of the initial first-pass fixation on a word, regardless of how many fixations the word received), gaze duration (the sum of all fixations on the word before the eyes left it for the first time), and total reading time (the sum of all fixations made on the target word); see Table 3.3. Participants and items were entered as crossed random effects. A full random structure was initially specified for participants and items, to avoid being anti-conservative (Barr et al., 2013); the random structure was trimmed until the models converged. Effects were considered significant when, initially, $|t| > 1.96$.

In all of the lme models there were significant group differences: children displayed significantly longer first fixations, gaze durations, and total reading times than the adults (see Table 3.3). We focus upon significant effects of the experimental manipulations, and any interactions with participant group.[9]

---

[9] In Appendix B; B.1, skipping rates are provided in Table B.1.1. No generalized linear mixed models would converge for this measure. In addition, separate analyses were also undertaken for the adults and the children with regards to Model 1, as shown in Table B.1.2.

Table 3.3

Mean and Standard Deviation (in parentheses) Reading Times on the Target Word in Each Condition

Group	Condition	First fixation duration (ms)	Gaze duration (ms)	Total reading time (ms)
Adults	123456	220 (66)	245 (84)	330 (186)
	ddd456	255 (89)	293 (109)	415 (235)
	1ddd56	254 (90)	291 (121)	395 (243)
	12ddd6	246 (78)	285 (99)	390 (224)
	123ddd	265 (98)	304 (118)	424 (276)
	dddddd	259 (79)	310 (131)	427 (242)
Children	123456	290 (141)	505 (509)	726 (672)
	ddd456	320 (159)	529 (368)	790 (591)
	1ddd56	293 (130)	505 (344)	773 (613)
	12ddd6	294 (144)	515 (388)	733 (511)
	123ddd	298 (162)	580 (595)	851 (746)
	dddddd	309 (155)	529 (476)	775 (595)

3.4.1 Model 1

This model used the identity control condition (123456) as a baseline, with each of the nonword preview conditions compared to it, thus examining the potential costs associated with substitutions being present in the parafovea, and the extent to which participants were gaining preview benefit. As can be seen from Tables 3.3 and 3.4, for all of the nonword preview conditions the adults experienced a significant cost relative to the identity condition- their foveal word identification was facilitated by obtaining a processing benefit from the correct parafoveal preview. The presence of significant interactions with participant group suggests that adults and children differed in their processing of letter substitutions in preview, in the earlier measure of first fixation duration. In contrast to the adults, children showed little increase in reading times for any of the substitution conditions, with the exception of ddd456; demonstrating a lack of preview benefit. Clearly, both adults and children, though, experienced a cost to early measures of lexical processing when parafoveal pre-processing of the first letter of the word was disrupted. Substitutions of other letters in the word disrupted very early lexical processing for adults but not children, who showed delayed sensitivity to substitutions of all except the first letter of the word. Certainly by the time the reader had engaged in second-pass reading on a word, both adults and children showed a cost to reading times from substitutions in all letter positions in preview, demonstrating comparable preview benefit effects.[10]

---

[10] Note that in Table 3.3, the means suggests that the children were not necessarily patterning in the same way as adults, especially in regard to gaze durations. It is also clear that there was substantially more variability around the means in the children's data compared to that of the adults. Indeed, within the separate analyses for the children (see Appendix B; B.1; Table B.1.2), after controlling for multiple comparisons, the two middle internal letter substitution preview conditions became non-significant; suggesting that, in gaze duration, the external letters being preserved in preview was as facilitative to the children's lexical identification as the identity condition (whilst the adults experienced significant costs).

Table 3.4

Output from Model 1 for First Fixation Duration, Gaze Duration, and Total Reading Time

	First fixation duration				Gaze duration				Total reading time			
	b	SE	*t*	*p*	b	SE	*t*	*p*	b	SE	*t*	*p*
Adults, 123456 (Int)	5.35	.03	184.63	< .001	5.44	.05	113.14	< .001	5.67	.06	101.81	< .001
Adults, Children	**.22**	**.04**	**5.34**	**< .001**	**.52**	**.06**	**8.04**	**< .001**	**.65**	**.07**	**8.84**	**< .001**
Adults, ddd456	**.13**	**.03**	**4.72**	**< .001**	**.17**	**.03**	**5.27**	**< .001**	**.23**	**.04**	**6.20**	**< .001**
Adults, 1ddd56	**.13**	**.03**	**4.73**	**< .001**	**.16**	**.03**	**4.93**	**< .001**	**.18**	**.04**	**4.76**	**< .001**
Adults, 12ddd6	**.10**	**.03**	**3.59**	**< .001**	**.15**	**.03**	**4.61**	**< .001**	**.17**	**.04**	**4.57**	**< .001**
Adults, 123ddd	**.17**	**.03**	**6.02**	**< .001**	**.21**	**.03**	**6.47**	**< .001**	**.23**	**.04**	**6.03**	**< .001**
Adults, dddddd	**.16**	**.03**	**5.69**	**< .001**	**.23**	**.03**	**6.94**	**< .001**	**.26**	**.04**	**6.95**	**< .001**
Children × ddd456	-.05	.04	-1.23	.220	-.06	.05	-1.25	.211	-.08	.05	-1.55	.121
Children × 1ddd56	**-.12**	**.04**	**-2.90**	**.004**	-.08	.05	-1.61	.108	-.05	.05	-.97	.331
Children × 12ddd6	**-.09**	**.04**	**-2.37**	**.018***	-.07	.05	-1.44	.151	-.07	.05	-1.38	.167
Children × 123ddd	**-.16**	**.04**	**-3.91**	**< .001**	-.09	.05	-1.80	.072	-.05	.05	-1.02	.309
Children × dddddd	**-.11**	**.04**	**-2.65**	**.008***	**-.15**	**.05**	**-3.27**	**.001**	**.13**	**.05**	**-2.51**	**.012***

Note. The reading time data were log transformed prior to analysis, so the model estimates cannot be directly interpreted. Significant effects are marked in bold. The syntax, following trimming, for first fixation duration, gaze duration, and total reading time as intercepts only models was as follows: *depvar ~ Group * condition + (1|Participant) + (1|targetno)*. The *s denote where significance levels changed with the use of the *glht* function (i.e., where results went from being significant to non-significant/marginally significant- within first fixation duration: *p* = .131 and *p* = .065, respectively, and within total reading time: *p* = .093).

3.4.2 Model 2

This model collapsed ddd456 and 123ddd together, and 1ddd56 and 12ddd6 together, in order to compare external to internal letter manipulations. The *contr.sdif* function (package MASS) was used to set up the factors. Then, contrasts were run to compare ddd456 to 123ddd for adults and children separately. As shown in Table 3.5, and Figure 3.3, the internal letter substitution conditions led to significantly faster reading times than the external letter substitution conditions, for both adults and children. Also, the contrasts revealed that, in first fixation duration, the children were showing a first-letter bias. Children's reading times were significantly slower in ddd456 than 123ddd in this very early measure of processing (see Table 3.3). Interestingly, note that in gaze duration and total reading time this effect of external letter substitutions seemed mainly to be driven by the end letter (123ddd; see Table 3.3).

Table 3.5

Output from Model 2, and Contrasts, for First Fixation Duration, Gaze Duration, and Total Reading Time

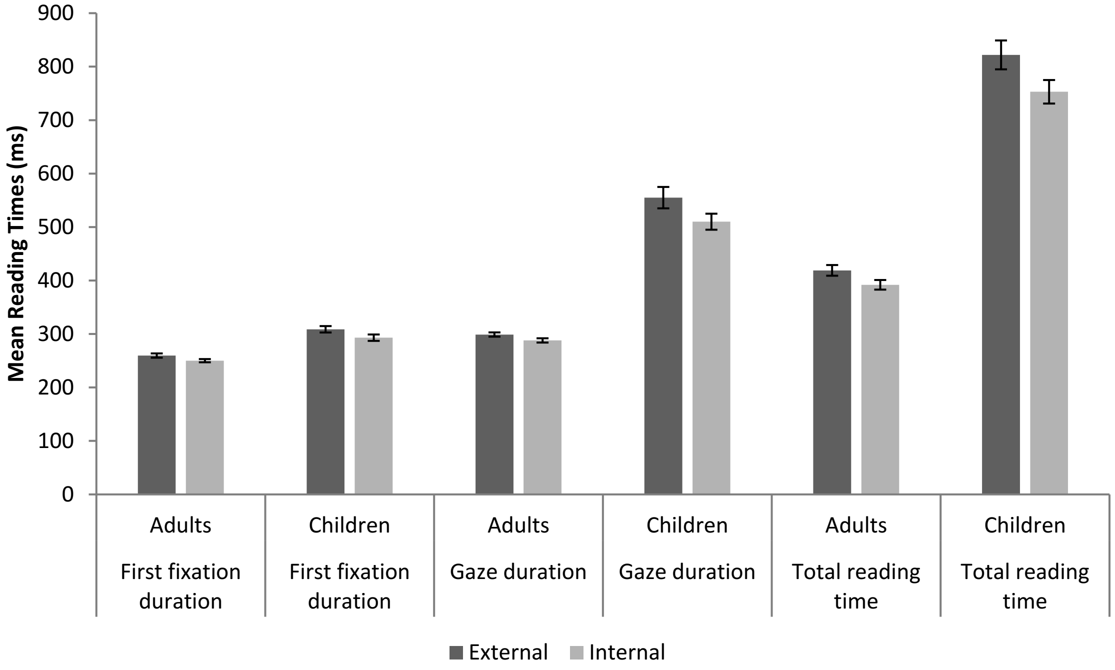
	First fixation duration				Gaze duration				Total reading time			
	b	SE	*t*	*p*	b	SE	*t*	*p*	b	SE	*t*	*p*
Intercept	5.54	.02	322.03	< .001	5.84	.03	183.00	< .001	6.16	.04	162.44	< .001
Adults, Children	**.11**	**.03**	**3.39**	**.001**	**.45**	**.06**	**8.07**	**< .001**	**.59**	**.07**	**8.99**	**< .001**
External vs. Internal	**-.03**	**.01**	**-2.44**	**.015**	**-.04**	**.02**	**-2.25**	**.024***	**-.05**	**.02**	**-2.92**	**.004**
Group × External vs. Internal	-.0003	.03	-.01	.991	.001	.03	.03	.978	.005	.04	.13	.894
Contrasts												
Intercept	5.53	.02	314.32	< .001	5.82	.04	144.40	< .001	6.14	.05	124.05	< .001
Adults, ddd456 - 123ddd	-.04	.03	-1.31	.189	-.04	.03	-1.22	.223	.006	.04	.17	.868
Children, ddd456 - 123ddd	**.07**	**.03**	**2.34**	**.019**	-.02	.03	-.57	.568	-.03	.04	-.74	.461

Note. The reading time data were log transformed prior to analysis, so the model estimates cannot be directly interpreted. Significant effects are marked in bold. The syntax for first fixation duration, gaze duration, and total reading time following trimming, as intercepts only models, was as follows: *depvar ~ Group * CollCons + (1|Participant) + (1|targetno)*. The contrasts were set up for first fixation duration, gaze duration, and total reading time within the following syntax (intercepts only models following trimming): *depvar ~ GroupByCond + (1|Participant) + (1|targetno)*. In order to use the *glht* function for Model 2, contrasts were set up for all dependent measures within the following syntax: *depvar*

~ *Group * condition3 + (1|Participant) + (1|targetno)*. The * denotes where the significance level changed with the use of the *glht* function (i.e., where the result went from being significant to marginally significant- $p = .071$).

Figure 3.3

Mean Reading Times for the Collapsed External Letter Substitution Conditions (ddd456 and 123ddd) and the Internal Letter Substitution Conditions (1ddd56 and 12ddd6), for Both Adults and Children.



3.4.3 Controlling for Multiple Comparisons

Given that Models 1 and 2 contain a number of comparisons across the five experimental conditions, we ran these models again using the *glht* function (package multcomp) to adjust *p* values and control for the multiple comparisons being made within each model (Hothorn et al., 2008).[11] For the majority of effects, this did not change the pattern of significance; we report here those instances where the correction did make a difference. First, within first fixation duration in Model 1, the interaction term between children and 12ddd6 became non-significant (and marginally significant between children

---

[11] We did not include the intercept when using the *glht* function, as it was not actively being compared within our models.

and dddddd), suggesting that the children's parafoveal pre-processing in these conditions was not significantly different (or only marginally so) to that of the adults.[12] Second, the interaction between children and dddddd became non-significant in total reading time; here, the children's processing was consistent with that of the adults (see also Appendix B; B.1; Table B.1.2). Third, within Model 2, in gaze duration, the main effect of external compared to internal letter substitutions in preview became marginally significant.

3.5 Discussion

The present study investigated parafoveal pre-processing in English children and adults during silent sentence reading, specifically comparing pre-processing of beginning, internal, and end letters. As expected, the children did pre-process the whole target word in the parafovea. Like adults, they displayed a cost from 123ddd substitutions, demonstrating that they were sensitive to substitutions of the final letter of the target words (albeit a slightly delayed effect, i.e., present in gaze duration). This indicates that children's parafoveal pre-processing (of word *n*+1) was not constrained by visual acuity limitations. If pre-processing was constrained by visual acuity, 123ddd should have been the least disruptive condition, as those substitutions were furthest away from the point of fixation. Instead, the significant cost associated with end letter substitutions clearly demonstrates that children's parafoveal pre-processing extended over the orthographic form of the whole word (six letters, in this case), rather than being constrained to the first few letters.

The data is suggestive of children's processing being delayed compared to the skilled adult readers, with a developmental change in the time course of pre-processing:

---

[12] When examining the children's first fixation duration results separately though (see Appendix B; B.1; Table B.1.2), the children's processing in these conditions was different to that of the adults: whilst the adults were showing costs in all of the preview conditions compared to the identity preview, the children were not (apart from marginally in the beginning letter substitution preview condition- ddd456).

adults showed early effects in first fixation duration, whilst the two groups only patterned similarly in later processing. This is consistent with children's rate of lexical processing being slower than that of adults, as found by the E-Z Reader model when used to simulate adults' and children's eye movement behaviour during reading (Reichle et al., 2013). If children are slower to process word *n* then it stands to reason that they will also be slower to pre-process information from word *n*+1. Consequently, each word in the sentence is pre-processed to a reduced degree, and is processed at a slower rate during direct fixation for a child compared to an adult. It is, therefore, unsurprising that children's overall reading times on words were longer, and that effects were delayed in children compared to adults.

This study provides strong evidence for the importance of external letters in children's lexical identification, consistent with skilled adult readers. As shown by collapsing, respectively, the internal and the external letter substitutions together, both adults and children benefitted from faster reading times when the internal letters (1ddd56 and 12ddd6), relative to the external letters (ddd456 and 123ddd), were substituted. Thus, consistent with the literature on skilled adult reading (White et al., 2008), the identity of a word's external letters facilitated children's parafoveal pre-processing more than its internal letters. With respect to syllabic boundaries, the conditions that substituted letters in both syllables of a word (1ddd56 and 12ddd6) were less disruptive to pre-processing than conditions that substituted letters in just one syllable (ddd456 and 123ddd). Thus, external letters are critical to parafoveal pre-processing, to a far greater degree than any pre-processing of syllabic structure.

These results are consistent with Grainger and Ziegler's (2011) model of orthographic processing. Both the adults and the children, albeit delayed, appeared to be using coarse-grained orthographic processing.[13] The benefits gained from the internal letter substitutions, relative to the external letter substitutions, suggests that both groups were not sensitive to the absolute precise ordering of letters in preview, but were rather coding for the most visible letters that best constrained word identity and facilitated lexical identification- the external letters. This is broadly supportive of flexible letter position encoding models (e.g., SOLAR, Davis, 2010a; SERIOL, Whitney, 2001).

The delay in the children's pre-processing of orthography (preview benefit) compared to the adults could be due to orthographic representations being less precisely encoded in the children (e.g., Perfetti, 2007). When letter substitutions were present in preview this came at an immediate cost to the adults compared to the identity condition, whilst this effect was delayed in the children. If orthographic forms are less precisely encoded in children, they would experience less of an immediate cost when orthography is manipulated in preview, in contrast to the adults with their more precisely encoded orthographic representations, who would be more reliant on the presence of whole-word orthography in preview (as provided by the identity condition). Consequently, there would appear to be a developmental change in the tuning of orthographic word recognition processes (e.g., Castles et al., 2007).

One unexpected result was the lack of a first-letter bias in the adults, that is a more important role in preview for the first letter than the final letter, as found in previous studies (e.g., White et al., 2008), though when first and final letters were collapsed into a single, "external" condition this was significantly different to internal letter substitutions

---

[13] This type of orthographic processing allows direct access from orthography to semantics (meaning). Using this kind of orthographic processing, it is posited that approximate letter positions are coded for within words. This is in contrast to fine-grained orthographic processing, which provides access to semantics via phonological and morphological forms, where sensitivity is shown to the precise order of letters within words.

(consistent with previous research). The present study did ultimately find though that the first letter of the target words was important to adults' pre-processing (albeit not more so than the final letter); substituting the first letters in preview (ddd456) came at a significant cost relative to the identity condition. It may be that the finding of a first-letter bias depends on the exact nature of the experimental manipulation. Most research has looked at letter transpositions, not substitutions (e.g., Johnson & Eisler, 2012; Rayner, White, et al., 2006; White et al., 2008). Importantly, though, Johnson et al. (Experiment 3; 2007), showed that both first letter transposition and substitution previews were detrimental to reading times.[14] Consequently, we would have expected an effect of first letter substitutions in the adults. The lack of this effect could be due to the stimuli which, here, were specifically designed for children and would, therefore, have been very easy for the skilled adult readers. The adults' ease of processing for these sentences may have resulted in a greater degree of parafoveal pre-processing for the target word than would be the case with more difficult sentences (e.g., Henderson & Ferreira, 1990). Thus, the adult readers may have allocated their attention across the entire form of word *n*+1 (not just the initial letters). For the adults, consequently, both the first and final (external) letters were important to their pre-processing.

Children, similar to the adults, displayed sensitivity to first letter substitutions very early in their lexical processing- in first fixation duration. The 30 ms preview benefit effect found within this measure in the children was comparable in size to the effect found within the adults (35 ms). This suggests that the privileged status of the first letter/s to lexical identification is evident very early in both adults' and children's lexical processing,

---

[14] Although no direct comparison was made of the first and final letter substitution conditions, differences in condition means suggest that the first letter substitution condition increased reading times more than the final letter substitution condition (Table 4, p. 218).

especially given how this information was manipulated parafoveally. Whilst the adults, though, did not show a first-letter bias (comparing ddd456 against 123ddd), the children did. This evidence for the importance of the first letter in children's pre-processing is consistent with Pagán et al. (2016) and Johnson et al. (2018), who found numerical trends for a bias towards the first bigram of target words in all dependent measures for children.[15] Overall, the evidence strongly suggests that the first letter/s of words are important for facilitating children's lexical identification in preview.

There are several reasons why the first letter of a word might be particularly important for lexical identification. One possibility is reduced lateral masking, or

---

[15] It is of note that the analyses undertaken by both Pagán et al. (2016) and Johnson et al. (2018) and the present study are, again, different. Pagán et al.'s study focused on comparing transposed letters to substituted letters (SLs), whilst the present study only examined substituted letters. Also, the present study included a final letter manipulation, Pagán et al.'s study did not. Johnson et al., similar to the present study, used letter substitutions in preview; however, although a final letter manipulation was present in this study, no direct comparison was, or could be, made with regard to its role in preview in comparison to the first letter, given that the final letter was manipulated in both orthographic preview conditions. Consequently, the closest comparison we could make to that of Pagán et al.'s SL12 versus SL23 effect is a comparison of ddd456 versus 1ddd56. We also show a numerical pattern in our dependent measures between ddd456 and 1ddd56 (first fixation duration: 27 ms; gaze duration: 24 ms; total reading time: 17 ms); these effects are larger than the largest effect found by Pagán et al. (10 ms in single fixation duration). It is likely that this is due to the different number of letters substituted; whilst Pagán et al. substituted two letters, we substituted three. The size of the effect is almost certain to have increased commensurately with the number of letters substituted. With regards to Johnson et al., the closest comparison we could make to their orthographically dissimilar preview versus their orthographically similar preview effect is dddddd versus 12ddd6. We also show a numerical pattern in our measures between dddddd and 12ddd6 (first fixation duration: 15 ms; gaze duration: 14 ms; total reading time: 42 ms), broadly consistent with their findings for the neutral context, as these results are most applicable to the present research (first fixation duration: 30 ms; gaze duration: 15 ms; total reading time: 19 ms).

crowding, due to the inter-word space on one side, whilst internal letters are subject to greater lateral masking from the presence of other letters on both sides (e.g., Bouma, 1973; Levi, 2008). Alternatively, it could be more cognitively based, in that identification of the first letter of a word could drive the process of lexical identification. Certainly, Johnson and Eisler's (2012) research, with adults, suggests this could be the case. For example, they found that when lateral masking was equated by replacing inter-word spaces with #s (e.g., The#boy#could#not#solve#the#problem#so#he#asked#for#help.), first letter transpositions were still significantly more difficult for readers than internal transpositions, whilst final letter transpositions were no more harmful than the internal transpositions (Experiments 1 and 2). This suggests a critically important role for the first letter of a word in lexical identification, irrespective of low-level visual factors like crowding. This finding contrasts with effects associated with a word's final letter.

In summary, the present study provides novel evidence of children pre-processing whole words during English reading, and experiencing costs from external letter manipulations in preview, similar to adults. External letters appear to play a specific and important role in visual word recognition, seeming to fundamentally relate to how both adult and child readers access lexical information.

Chapter 4 Phonological Parafoveal Pre-processing in Children Reading English Sentences

The data that supports the findings of this study, and the code used for the main model analyses, are available from:

https://osf.io/pzfhg/?view_only=ec4821493476463aa7cafecefe04dff3.

4.1 Abstract

Although previous research has shown that, in English, both adult and teenage readers parafoveally pre-process phonological information during silent reading, to date, no research has been conducted to investigate such processing in children. Here we used the boundary paradigm during silent sentence reading, to ascertain whether typically developing English children, like adults, parafoveally process words phonologically. Participants' eye movements were recorded as they read sentences which contained, in preview, correctly spelled words (e.g., *cheese*), pseudohomophones (e.g., *cheeze*), or spelling controls (e.g., *cheene*). The orthographic similarity of the target words available in preview was also manipulated to be similar (e.g., *cheese*/*cheeze*/*cheene*) or dissimilar (e.g., *queen*/*kween*/*treen*). The results indicate that orthographic similarity facilitated both adults' and children's pre-processing. Moreover, children parafoveally pre-processed words phonologically very early in processing. The children demonstrated a pseudohomophone advantage from preview that was broadly similar to the effect displayed by the adults, although the orthographic similarity of the pseudohomophone previews was more important for the children than the adults. Overall, these results provide strong evidence for phonological recoding during silent English sentence reading in 8- to 9-year-old children.

4.2 Introduction

Phonology (the pattern of speech sounds within a language) plays a key role in children's literacy acquisition (e.g., Share, 1995). Typically, learning to speak precedes learning to read; it is during learning to read that orthography (words' printed forms) is associated with pre-existing cognitive lexical entries that contain both phonological and semantic (meaning) information (Frost, 1998). It is widely accepted that processing of phonology is a critical component of learning to read. Phonological decoding (the overt, effortful, sounding out of letter sounds) is acknowledged as a vital early phase of reading acquisition (e.g., Ehri, 2007; Frost, 1998; Share, 1995). A pervasive question, therefore, within reading research is the extent to which phonology plays a role in word (lexical) identification, and how this may change through the development from beginning to skilled reader (for a recent review see Milledge & Blythe, 2019). Whilst it is known that phonology is important for children learning to read, much less is known about what happens when they become skilled enough to read silently and independently. In the present study, we examined the extent to which beginner 8- to 9-year-old readers of English were able to process phonological cues from an upcoming word during silent sentence reading, in comparison to skilled English adult readers.

Phonology clearly plays an important role in skilled adult readers' lexical identification processes; such readers have been shown to undertake phonological recoding (the covert, rapid pre-lexical processing of a word's phonology, that is, phonology becoming activated during lexical identification) (Leinenger, 2014). The clearest evidence for pre-lexical phonological processing in adults comes from research investigating parafoveal pre-processing. Adult readers do not only process the word they are directly fixating (word *n*) but also begin to process some information about the upcoming word (word *n*+1) prior to direct fixation (e.g., Rayner, 1998, 2009). This means that when word *n*+1 is eventually fixated, reading times are faster due to the processing the reader has already undertaken in relation to that word (see Schotter et al., 2012 for a review). This is referred to as parafoveal pre-processing, and it is typically examined using the boundary paradigm (Rayner, 1975). Using this paradigm researchers put an invisible boundary in the space immediately before a target word. Before the readers' eyes cross this boundary, a preview letter string is present in place of the target word. As soon as the readers' eyes cross this boundary the preview letter string changes to the correct target word. This paradigm allows researchers to manipulate the characteristics of the preview letter string in relation to the target word to examine the types of information that readers are able to process in the parafovea, and how such processing facilitates subsequent lexical identification. Faster reading times on the target word following a related preview compared to an unrelated preview letter string is referred to as preview benefit.

Research using this paradigm has shown that during silent sentence reading, skilled adult readers use phonological codes to aid lexical identification of word *n*+1; they phonologically process the upcoming word (Leinenger, 2014; Milledge & Blythe, 2019; Vasilev et al., 2019). For example, both Chace et al. (2005) and Pollatsek et al. (Experiment 2; 1992) found that adult readers displayed faster reading times on a correct target word when a homophone (e.g., *beech* as a preview for *beach*) was present in preview before the readers' eyes crossed the boundary, compared to a spelling control preview (e.g., *bench* as a preview for *beach*). Interestingly, recent research has shown that such effects (i.e., faster reading times on words due to phonological and orthographic preview similarity, compared to previews with no orthographic or phonological overlap) are even evident cross-script in Russian-English bilinguals (Jouravlev & Jared, 2018).

Moreover, Pollatsek et al. (1992) also found that orthographic similarity of the preview to the correct target word affected participants' pre-processing of phonology: the greater the orthographic overlap between the correct target word and its homophone preview, then the shorter the reading times on the target word (e.g., *paste* as a preview for *paced* resulted in faster reading times than *shoot* as a preview for *chute*). Pollatsek et al. posited, therefore, that phonological and orthographic (graphemic) codes jointly aid lexical identification of an upcoming word. Thus, the facilitatory effects they observed were a function of both the orthographic and phonological overlap of the previews, suggestive of an interactive relationship between orthography and phonology within parafoveal pre-processing. This interactive effect has also been reported in several other studies (e.g., Blythe et al., 2018, 2020). What exactly drives this interactive relationship, though, has never fully been explained.

We suggest that the importance of the first letter of an upcoming word in preview could be critical in respect to why phonological pre-processing is modulated by orthographic similarity. Past research has shown that the first letter plays a vital role in skilled adult readers' ability to lexically identify a word, both under direct fixation and, critically, during parafoveal pre-processing (e.g., Briihl & Inhoff, 1995; Inhoff, 1987, 1989a,b; Johnson & Eisler, 2012; Johnson et al., 2007; Milledge et al., 2020; White et al., 2008). In the studies that report an interaction between orthography and phonology in preview, it is typically shown that the advantage of having phonological information preserved in preview is greater when the orthographic manipulation involves fewer letter substitutions (e.g., *foul – fowl – foil*) than when it involves more letter substitutions (e.g., *chute – shoot – shout*). As this example shows, the number of letter substitutions is carefully controlled across the homophone and the spelling control (e.g., four out of five letters, in the same within-word locations, are substituted to form both *shoot* and *shout* from the target word *chute*). The manipulations designated as orthographically dissimilar typically involve both more letter substitutions and, very importantly in relation to our suggestion, often substitute the first letter of the word. For example, Pollatsek et al. (1992) found that, within first fixation duration, on average, adults did not show as much benefit from phonology (a homophone; e.g., *shoot* as a preview for *chute*), over a spelling control (e.g., *shout* as a preview for *chute*), when the first letter was substituted in preview (e.g., first letter *c* substituted with *s* in preview; *shoot – chute*) compared to when the first letter was maintained in preview (e.g., *fowl* as a preview for *foul* vs. *foil*) (20 ms benefit vs. 37 ms benefit, respectively). In addition, for homophone previews that maintained the first letter, first fixation durations, on average, were more comparable with the identity previews (11 ms cost), in comparison to the homophone previews that substituted the first letter (24 ms cost). We suggest that preserving the orthographic code of the first letter of a word in preview facilitates the extraction of phonological information from the parafovea and lexical identification of a word in adult readers.

Research has also shown that typically developing teenagers are able to pre-process phonology (Blythe et al., Experiment 2; 2018, 2020). Within both studies, teenagers displayed evidence of a pseudohomophone advantage (faster reading times on a pseudohomophone preview compared to a spelling control preview). Moreover, similar to the results of Pollatsek et al. (1992), it was also found that orthographic similarity played a role in the teenagers' pre-processing: previews that were orthographically similar to the correct target word (e.g., *cherch/charch* as previews for *church*) resulted in faster reading times than orthographically dissimilar previews (e.g., *kween/treen* as previews for *queen*). Phonology clearly plays an important role in pre-lexical processing during silent sentence reading, albeit contingent on the degree of orthographic similarity, facilitating both adult and teenage readers' pre-processing of word *n*+1.

As yet, no eye movement research has examined whether beginner child readers of English also pre-lexically process phonology in the parafovea, similar to skilled adult and teenage readers. Two pieces of research, though, have shown evidence of children processing phonology during direct fixation. Blythe et al. (2015) provide evidence of this in children as young as 7-years-old. Blythe et al. recorded the eye movements of 7- to 9-year-old children, and adults, as they silently read sentences that contained a correct target word (e.g., *water*), a pseudohomophone (e.g., *worta*), or a spelling control (e.g., *wecho*). Both the adult and child participants displayed significantly faster reading times when a pseudohomophone was present compared to a spelling control; suggesting that the valid phonology of the pseudohomophone facilitated their lexical identification. Similarly, Jared et al. (Experiment 3; 2016) have also shown that 10- to 11-year-olds use phonological codes to access a word's lexical representation during silent sentence reading. The children read sentences that contained either a correct target word (e.g., *whether*), a homophone (e.g., *weather*), or a spelling control (e.g., *winter*). Critically, the children displayed faster reading times when the homophone was present in a sentence compared to the spelling control. During silent sentence reading, therefore, phonology appears to play a key role in children's lexical identification. In both of these studies, though, phonological processing occurred during direct fixation of the pseudo/homophones. All of this said, it remains unknown whether phonological processing plays a role in pre-lexical processing of a word in the parafovea in beginner readers of English.

Recent research, though, has shown that child readers of English are sensitive to orthographic information in the parafovea. Pagán et al. (2016) found that 8- to 9-year-old children were similarly affected to adults by letter substitutions and transpositions in preview, in regard to both time course and magnitude of effects. Similarly, Milledge et al. (2020) also found that 8- to 9-year-old children, like adults, were sensitive to letter substitutions in preview across the whole-word form (six letters), and, more specifically, external letter substitutions in preview (e.g., *savber*/*numtoc* as previews for *number*) were more harmful to both the adults' and the children's processing than internal letter substitutions in preview (e.g., *navter*/*nuvtor* as previews for *number*). Moreover, an early first-letter bias was found, such that reading times were longer when the first letter of a word was substituted in preview compared to when the end letter was substituted in preview. In addition, Johnson et al. (2018) found that child readers (6- to 12-year-olds) were sensitive to orthography in preview, as they displayed longer reading times when an orthographically dissimilar letter string was present in preview (e.g., *esium* as a preview for *ocean*) compared to an orthographically similar preview (e.g., *ocium* as a preview for *ocean*). Moreover, both of these previews came at a cost to processing relative to the identity condition (where no display change occurred). Overall, this research clearly demonstrates that, similar to skilled adult readers, children from the age of 8-years pre-process the orthography of word *n*+1 as an integral aspect of lexical processing during sentence reading, such that there is a cost to the efficiency of their lexical processing if such pre-processing is disrupted. Whilst, however, it is known that orthography plays a role in children's parafoveal pre-processing, it is unknown whether orthography modulates phonological pre-processing in children, as is the case for adults (Pollatsek et al., 1992).

There are a number of theories of word recognition that posit how both orthography and phonology might contribute to lexical identification (e.g., Coltheart et al., 2001; Grainger & Ziegler, 2011; Perry et al., 2007). These theories typically focus upon isolated word recognition that occurs when a word is under direct fixation. Whilst these theories do not directly offer any account of the role of parafoveal processing in lexical identification, they do still have the potential to provide insight into the nature of such processing as it occurs during natural reading. For example, Pagán et al. (2016) argued that the letter position encoding effects they observed in parafoveal preview during sentence reading were consistent with the theory offered by Grainger and Ziegler (2011). Grainger and Ziegler (2011) proposed that both phonological and orthographic characteristics of lexical stimuli exert an influence in lexical identification via two processing routes: a coarse-grained processing route and a fine-grained processing route. The coarse-grained route allows semantic information to be directly accessed from orthographic form and permits some flexibility with regard to orthographic encoding (i.e., misspellings can be tolerated). The fine-grained route allows access to semantics via commonly co-occurring letter patterns being processed and mapped onto their corresponding phonological representations. Within this route, though, there is little flexibility with regard to orthographic encoding- the first letter's correct orthographic code being present could be especially important to this processing, enabling efficient processing of phonological code(s). According to this theory, early in reading acquisition, children are not expected to show rapid, pre-lexical influences of phonology during word recognition (i.e., they would not display phonological recoding). However, as their age and reading skill increases, children should show a decrease in their reliance on phonological decoding (the slow, laborious, serial sounding out of letter sounds) and an increased reliance on coarse-grained processing, along with fine-grained processing that allows phonological recoding.[16] This fine-grained processing has been evidenced in adult and teen readers (Blythe et al., 2018, 2020; Pollatsek et al., 1992); both groups have been shown to pre-process phonology from word *n*+1 (phonological recoding) and, importantly, this was modulated by orthographic similarity. Orthographically dissimilar previews would be expected to disrupt processing within the fine-grained route due to the greater discrepancy between orthographic and phonological information (given the fine-grained route's limited tolerance for word misspellings; e.g., as at least two letters were substituted in preview and, importantly, half of the previews involved at least the first letter being substituted). In contrast, the orthographically similar previews would be expected to cause less disruption to processing within the fine-grained route as only one letter was manipulated in preview. Critically, within the orthographically similar previews the first letter was never substituted.

With regard to children, Grainger and Ziegler's (2011) theory raises the question of whether there is a point in reading development when phonological decoding is not used

---

[16] Note that what Grainger and Ziegler (2011) refer to as "phonological recoding", is referred to as phonological decoding within the present paper.

but rapid phonological processing is not yet fully developed and efficient, as it is for adult and teenage readers. If this is the case, it is possible that whilst children might display benefits from phonology under direct fixation (Blythe et al., 2015; Jared et al., 2016), they might not be able to process phonology as rapidly as is required within parafoveal pre-processing (i.e., 8- to 9-year-olds might not be able to extract phonological codes from word *n*+1).

In the present study, we examined parafoveal pre-processing of phonology and orthography in a typical population of 8- to 9-year-old native readers of English, in comparison to a group of skilled adult readers. We manipulated the orthographic and phonological features of the target words in preview, allowing us to examine: (1) whether beginner readers of English pre-process phonology; (2) whether phonological and orthographic pre-processing might be independent or interactive (if they are interactive it would be supportive of readers of English using the fine-grained route of processing with regard to processing of word *n*+1; Grainger & Ziegler, 2011); and (3) whether the pre-processing undertaken by participants differed between skilled adult and beginner child readers.

The boundary paradigm (Rayner, 1975) was used to examine these questions, with two manipulations made: (1) phonological similarity; and (2) orthographic similarity. Participants either had in preview the correct target word (identity; e.g., *cheese*), a pseudohomophone (e.g., *cheeze*), or a spelling control (e.g., *cheene*). In addition, the orthographic similarity of the previews to the correct target word was also manipulated (e.g., *cheeze/cheene* as orthographically similar previews for *cheese* vs. *kween/treen* as orthographically dissimilar previews for *queen*). Regarding the three research questions, (1) it was unknown what role, if any, phonology would play in preview for the beginner child readers, as no research, thus far, has examined the role such information being present might play for beginner readers of English when considering word *n*+1. (2) It was predicted that, for the adult readers, phonological and orthographic pre-processing were likely to be interactive given past research findings (e.g., Pollatsek et al., 1992), that is, adult readers would display a pseudohomophone advantage (faster reading times on the target word after a pseudohomophone preview compared to a spelling control preview) and that this advantage would be modulated by orthographic similarity (i.e., present within the orthographically similar previews, where the first letter of a target word was maintained, but absent within the orthographically dissimilar previews, given that the first letter of a target word was substituted in half of these previews). Given that child readers of English have been found to pre-process whole-word orthography (Johnson et al., 2018; Milledge et al., 2020; Pagán et al., 2016), it was predicted that they would display faster reading times on the target words after orthographically similar previews compared to orthographically dissimilar previews; whether there would be an interactive effect with phonology, as predicted for the adults (e.g., Pollatsek et al., 1992), was unknown. Milledge et al.'s (2020) finding, though, of a first-letter bias within 8- to 9-year-old children could suggest that an interactive effect could be present. Obtaining interactive effects within the 8- to 9-year-old child readers would not only be indicative of phonological recoding being undertaken, but also of fairly sophisticated parafoveal pre-processing occurring whereby phonological codes are accessed from orthographic codes via a fine-grained route of processing (Grainger & Ziegler, 2011). (3) Apart from predicted overall group differences (i.e., the children displaying longer reading times than the adults; e.g., Blythe & Joseph, 2011), it was unknown whether the children would differ in their pre-processing to the adults.

4.3 Method

4.3.1 Participants

Forty-eight adults (*M* = 21.02 years old; *SD* = 3.56) from the University of Southampton community and 48 children (aged 8- to 9-years-old; *M* = 8.31; *SD* = .47) from a local junior school participated in the eye-tracking experiment. See Table 4.1 for a summary of group characteristics. All participants had normal or corrected to normal vision, and were native speakers of English with no known reading difficulties, as confirmed by the reading subtests of the Wechsler Individual Achievement Test II UK (WIAT-II UK; Wechsler, 2005). All participants' composite standardised scores were within the expected range (adults' score range: 99-139; children's score range: 99-136; see also Table 4.1). All adult participants gave informed written consent prior to participation. Parents provided informed written consent on behalf of their children, and the children also provided their own informed written assent, prior to participation. Ethical approval was provided by the University of Southampton Psychology Ethics Committee (submission ID: 45888).

We recruited a larger number of participants than Blythe et al. (2018, 2020), who had 23 participants in each group (2018) and 30 older typically developing teenage readers (2020). Regarding the number of stimuli, outlined in the subsection below, the English language was exhausted in order to create these tightly controlled stimuli.

Table 4.1

Summary of Group Characteristics

		Mean	StDev	*t*	*df*	*p*
Test age (years)	Adults	21.02	3.56			
	Children	8.31	.47	24.49	94	< .001
WIAT word reading	Adults	113.77	5.41			
	Children	112.79	6.58	.80	94	.428
WIAT pseudoword decoding	Adults	108.21	6.50			
	Children	108.27	5.67	- .05	94	.960
WIAT comprehension	Adults	116.29	6.08			
	Children	114.21	6.72	1.59	94	.114
WIAT composite standardised scores	Adults	119.35	9.78			
	Children	113.35	8.32	3.24	94	.002

Note. The three right-hand columns give the results of independent samples *t*-tests comparing the adults to the children. All WIAT scores are standardised. The significant difference in the composite scores indicates that the children were, for their age, less skilled than the adult readers (perhaps unsurprisingly given the more heterogeneous sample within a state junior school compared to students within a university). Importantly, both participant groups were reading at or above the expected level, with no evidence of reading difficulties for any individual participant.

4.3.2 Materials and Design

We used the stimuli from Experiment 2 reported in Blythe et al. (2018, 2020), comprised of 24 target words and sentence frames, which had been pre-screened with 78 8-

to 9-year-old children. Two manipulations were made: a within-item phonological manipulation, and a between-item orthographic manipulation. Regarding the phonological manipulation, two nonwords were created for each target word to create a triplet of previews: the correctly spelled (identity) preview; a pseudohomophone preview; and a spelling control preview (e.g., *cheese*/*cheeze*/*cheene*). The length of the previews was always matched, and syllabic structure was maintained. The nonword previews had also been matched on orthographic overlap with the identity preview, number of orthographic neighbours, consonant-vowel structure, and word shape (e.g., descenders were substituted with descenders, ascenders with ascenders, etc.).

For the orthographic manipulation, each preview triplet was either orthographically similar (12 triplets) or orthographically dissimilar (12 triplets). Within the orthographically similar triplets, only one letter (never the first or second letter) was substituted to form the two nonword previews (e.g., *cheese*/*cheeze*/*cheene*). Within the orthographically dissimilar triplets, at least two letters (with one letter at least being the first and/or second letter) were substituted to form the two nonword previews (e.g., *queen*/*kween*/*treen*).

The correctly spelled (identity condition) target words (12 in each condition) were matched on frequency from an adult corpus (0-1882 per million; Balota et al., 2007) and a child corpus (8-560 per million; Masterson et al., 2003), Age of Acquisition (2.90 - 7.63 years; Kuperman et al., 2012), and orthographic neighbourhood size (0-23) (all *t*s < 2, all *p*s > .1). Both sets of stimuli contained 4-6 letter words (with a marginally significant difference in word length between the two sets of stimuli, $p = .05$).

Three counterbalanced lists of sentences were created, each including either an identity preview, a pseudohomophone preview, or a spelling control preview from each triplet: four of each kind of preview from the orthographically similar stimuli, and four from the orthographically dissimilar stimuli. Consequently, each participant read 24 sentences; containing eight identity previews, eight orthographically similar previews (four pseudohomophone previews and four spelling control previews), and eight orthographically dissimilar previews (four pseudohomophone previews and four spelling control previews). The sentences occupied one line on the screen (maximum = 70 characters; *M* = 61 characters; e.g., *Cheddar is my favourite kind of cheese to have for lunch.*). Example stimuli in each condition are shown in Figure 4.1.

Figure 4.1

Example experimental sentences showing the three parafoveal preview conditions, with the invisible boundary shown (though shown as a visible line here it was not visible to participants). The asterisks refer to a fixation, the target word is shown in bold, and the condition is shown in the brackets (Id- identity; PSH- pseudohomophone preview; SC- spelling control preview). The first sentence gives an example of orthographically similar previews and the second sentence gives an example of orthographically dissimilar previews.

1. Cheddar is my favourite kind of **cheese** to have for lunch. **(Id)**
2. Cheddar is my favourite kind of **cheeze** to have for lunch. **(PSH)**
3. Cheddar is my favourite kind of **cheene** to have for lunch. **(SC)**

*

1. People cheered for the king and **queen** as they waved from the window. **(Id)**
2. People cheered for the king and **kween** as they waved from the window. **(PSH)**
3. People cheered for the king and **treen** as they waved from the window. **(SC)**

*

4.3.3 Apparatus and Procedure

Participants first completed the eye-tracking experiment. An EyeLink 1000 eye-tracker recorded right eye movements (SR Research). Participants were seated comfortably using forehead and chin rests, to minimise head movements, and were instructed to read normally and for comprehension. Then a three-point horizontal calibration and validation procedure was carried out. If the mean validation error, or the errors for any of the individual points, was greater than .20°, then the procedure was repeated. There were four practice trials at the beginning of the experiment (with two comprehension question trials), to ensure that participants were familiar with the procedure. A single sentence was presented at a time in 14-point, black Courier New font on the grey background of a 21 in. CRT monitor, with a refresh rate of 120 Hz, at a 60 cm viewing distance; one character subtended .34° of visual angle. Once participants had finished reading a sentence, they pressed a button on a gamepad, and nine of the sentences were followed by a comprehension question, to which the participants responded. After completion of the experiment, participants were asked whether they had noticed anything strange about the appearance of the sentences in the experiment: detecting a display change can affect fixation times (e.g., White et al., 2005). Two adult participants reported noticing something unusual about the sentences, even though they could not specify what, so their data were excluded from the analyses. Then participants completed the three reading subtests of the WIAT-II UK (Wechsler, 2005). The whole experiment lasted about 35 minutes per participant.

4.4 Results

All participants scored at least 77% on the comprehension questions (adults: *M* = 97.69%, *SD* = 4.56%; children: *M* = 90.97%, *SD* = 9.07%). The data were trimmed using the clean function in DataViewer (SR Research).[17] In total 884 fixations were merged or

---

[17] Fixations less than 80 ms were merged with the neighbouring fixation if within a .50° distance of another fixation over 80 ms. Also, fixations less than 40 ms were merged with neighbouring fixations if within a 1.25° distance of each other. If an interest area had three or more fixations less than 140 ms, these were then

deleted (2.22% of the dataset; 385 adult fixations and 499 child fixations), resulting in a final dataset of 38,929 fixations.

Data were analysed using linear mixed effects (lme) models, using the *lmer* function from the lme4 package (Bates et al., 2015) within the R environment for Statistical Computing (R Core Team, 2020), with participants and items entered as crossed random effects. To avoid being anti-conservative, a full random structure was initially specified for participants and items (Barr et al., 2013). If the models for each dependent measure failed to converge, the random structure was trimmed until they did converge. Data (for both global and local analyses) were log transformed before analysis, to reduce skew.[18]

4.4.1 Global Measures

Firstly, we examined global measures of participants' eye movement behaviour- their eye movements across entire sentences. As can be seen from Table 4.2, the children displayed significantly longer total sentence reading times ($b = .53$, $SE = .06$, $t = 9.22$, $p < .001$), longer fixation durations ($b = .10$, $SE = .02$, $t = 4.69$, $p < .001$), and made more fixations ($b = .40$, $SE = .05$, $z = 8.54$, $p < .001$) than the adults, consistent with past research (e.g., Blythe et al., 2009, 2011; Blythe & Joseph, 2011; Joseph et al., 2009; Tiffin-Richards & Schroeder, 2015a,b).[19]

---

merged into longer fixations. Subsequently, all remaining fixations less than 80 ms or greater than 1,200 ms were deleted.

[18] Note that, within the global analyses, the fixation count data was not log transformed and was analysed using a generalized linear mixed model, so the Poisson distribution could be used.

[19] Following trimming, the syntax for fixation count was: *FIXATION_COUNT ~ Group + (1|Participant) + (1 + Group|targetno)*, the syntax for total sentence reading time was: *IP_DURATION ~ Group + (1|Participant) + (1 + Group|targetno)*, and the syntax for fixation duration, as an intercepts only model, was: *CURRENT_FIX_DURATION ~ Group + (1|Participant) + (1|targetno)*.

Table 4.2

Mean and Standard Deviation (in parentheses) Values for Measures across Entire Sentences

Measure	Adults	Children
Fixation duration (ms)	239 (120)	267 (141)
Fixation count	14 (4)	20 (7)
Total sentence reading time (ms)	3260 (1269)	5636 (2351)

4.4.2 Local Measures

Then, we analysed reading time data on the target word in each sentence. Before analysing the local dependent measures, the data were further cleaned: trials in which the boundary change occurred early during a fixation on the pretarget word, and those that occurred late when the display change was not completed until more than 15 ms after onset of fixation on the target word were excluded from the analyses (110 adult trials- 9.55% of the adult trials, and 140 children's trials- 12.15% of the children's trials).[20]

---

[20] We undertook two analyses relating to late boundary changes. In the first, a late change was categorised as one occurring 10 ms or later than it being triggered, and in the second, we used a 15 ms criterion. Due to the pattern of data remaining unchanged between the two sets of analyses, the 15 ms criterion was adopted as it allowed for the retention of more data (2,054 data points as opposed to 1,922). Regarding the number of items per condition for each participant, after the boundary change cleaning, the lowest total number of items recorded for an adult participant was 17 ($M = 21.71$, total range: 17-24; identity $M = 7.33$, range: 5-8; orthographically similar pseudohomophones $M = 3.69$, range: 1-4; orthographically dissimilar pseudohomophones $M = 3.46$, range: 2-4; orthographically similar spelling controls $M = 3.58$, range: 2-4; and orthographically dissimilar spelling controls $M = 3.65$, range: 2-4) and within the child participants this was

The standard, key dependent measures were: first fixation duration (the duration of the initial first-pass fixation on a word, regardless of how many fixations the word received), single fixation duration (the time that a word was fixated when it received only one first-pass fixation), and gaze duration (the sum of all fixations on the word before the eyes left it for the first time); see Table 4.3.[21]

---

also 17 (*M* = 21.08, total range: 17-23; identity *M* = 7.13, range: 5-8; orthographically similar pseudohomophones *M* = 3.35, range: 2-4; orthographically dissimilar pseudohomophones *M* = 3.50, range: 2-4; orthographically similar spelling controls *M* = 3.44, range: 2-4; and orthographically dissimilar spelling controls *M* = 3.67, range: 2-4).

[21] The probability of the children making a single fixation across all trials was .65 and the probability of the adults making a single fixation across all trials was .75. Single fixation probabilities for the adults and the children by condition are available in Appendix B; B.2 (Table B.2.1). Within Appendix B; B.2, skipping rates are also provided in Table B.2.2. Within the main model analyses, no generalized linear mixed models would converge for this measure. Intercept only models within the contrasts converged but the results were non-significant, *p*s > .311.

Table 4.3

Mean and Standard Deviation (in parentheses) Reading Times on the Target Word in Each Condition

Group	Condition	First fixation duration (ms)	Single fixation duration (ms)	Gaze duration (ms)
Adults	Identity	218 (81)	218 (83)	237 (106)
	Orthographically similar pseudohomophones	222 (71)	223 (71)	253 (102)
	Orthographically similar spelling controls	235 (75)	239 (71)	265 (97)
	Orthographically dissimilar pseudohomophones	236 (77)	249 (78)	269 (83)
	Orthographically dissimilar spelling controls	246 (77)	253 (81)	275 (92)
Children	Identity	247 (103)	258 (107)	317 (161)
	Orthographically similar pseudohomophones	252 (93)	267 (99)	329 (173)
	Orthographically similar spelling controls	275 (111)	294 (110)	341 (144)
	Orthographically dissimilar pseudohomophones	307 (128)	330 (128)	363 (139)
	Orthographically dissimilar spelling controls	289 (124)	315 (129)	361 (153)

Two lme models were run. Model 1 compared each of the nonword conditions to the correctly spelled identity preview condition, with participant group included as an interaction. This allowed us to look at the costs associated with a nonword being present in preview, examining whether participants displayed preview benefit, with the children being compared to the adults. Model 2 excluded the correctly spelled identity preview, and only included the nonword preview conditions, again with participant group included as an interaction. This model allowed us to directly examine the effects of phonology (pseudohomophones vs. spelling controls) and orthography (orthographically similar vs. orthographically dissimilar) on both adults' and children's parafoveal pre-processing.

This two-step approach was used due to the unbalanced nature of this experiment, given the between-items manipulation of orthographic similarity. Whilst the orthographic similarity/dissimilarity split was meaningful regarding the nonword (pseudohomophone and spelling control) previews, this split was not meaningful regarding the correctly spelled target word (identity) previews (i.e., one-third of the stimuli would have been incorrectly classed as orthographically similar or dissimilar). Thus, for Model 1 the data from the identity previews was collapsed into a single condition. Within the two models, effects were considered significant when $|t| > 1.96$.

4.4.2.1 Model 1

The five experimental conditions were: 1) identity previews; 2) orthographically similar pseudohomophone previews; 3) orthographically dissimilar pseudohomophone previews; 4) orthographically similar spelling control previews; and 5) orthographically dissimilar spelling control previews. The reading times for adults on the identity previews provided the intercept for this model. The results of this model, for each of the dependent measures, are shown in Table 4.4.

Table 4.4

Output from Model 1 for First Fixation Duration, Single Fixation Duration, and Gaze Duration

	First fixation duration				Single fixation duration				Gaze duration			
	b	SE	*t*	*p*	b	SE	*t*	*p*	b	SE	*t*	*p*
Adults, Identity (Int)	5.32	.03	184.85	< .001	5.32	.03	161.44	< .001	5.38	.04	147.50	< .001
Adults, Children	**.11**	**.04**	**3.16**	**.002**	**.17**	**.04**	**4.39**	**< .001**	**.27**	**.04**	**6.66**	**< .001**
Orthographically similar pseudohomophones	.05	.03	1.57	.117	.05	.03	1.59	.111	**.09**	**.04**	**2.43**	**.015**
Orthographically dissimilar pseudohomophones	**.09**	**.04**	**2.67**	**.008**	**.14**	**.03**	**4.08**	**< .001**	**.16**	**.04**	**4.52**	**< .001**
Orthographically similar spelling controls	**.10**	**.03**	**3.04**	**.002**	**.13**	**.03**	**3.85**	**< .001**	**.14**	**.04**	**3.77**	**< .001**
Orthographically dissimilar spelling controls	**.13**	**.03**	**4.08**	**< .001**	**.15**	**.03**	**4.31**	**< .001**	**.18**	**.04**	**4.85**	**< .001**
Children × Orthographically similar pseudohomophones	-.009	.04	-.20	.839	-.008	.05	-.17	.865	-.03	.05	-.69	.490
Children × Orthographically dissimilar pseudohomophones	**.12**	**.04**	**2.63**	**.009**	.08	.05	1.62	.106	-.004	.05	-.08	.940
Children × Orthographically similar spelling controls	.02	.04	.40	.686	.001	.05	.01	.989	-.03	.05	-.68	.497
Children × Orthographically dissimilar spelling controls	.02	.04	.36	.716	.04	.05	.78	.433	-.02	.05	-.38	.701

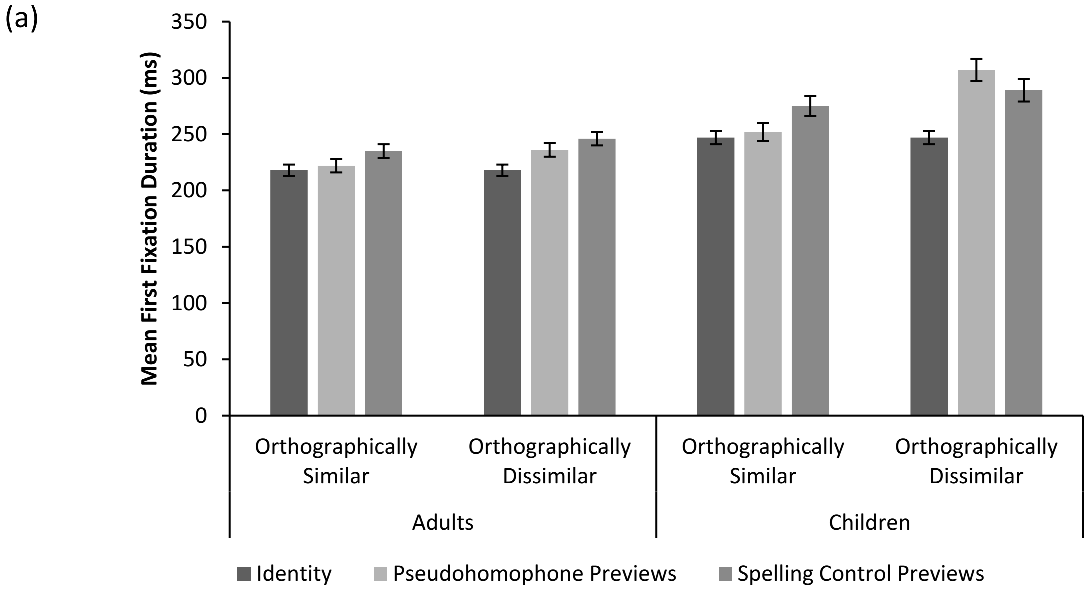
Note. The reading time data were log transformed prior to analysis, so the model estimates cannot be directly interpreted. Significant effects are

marked in bold. Following trimming, as intercepts only models, the syntax for all dependent measures was as follows: *depvar ~ Group * Condition + (1|Participant) + (1|targetno)*.

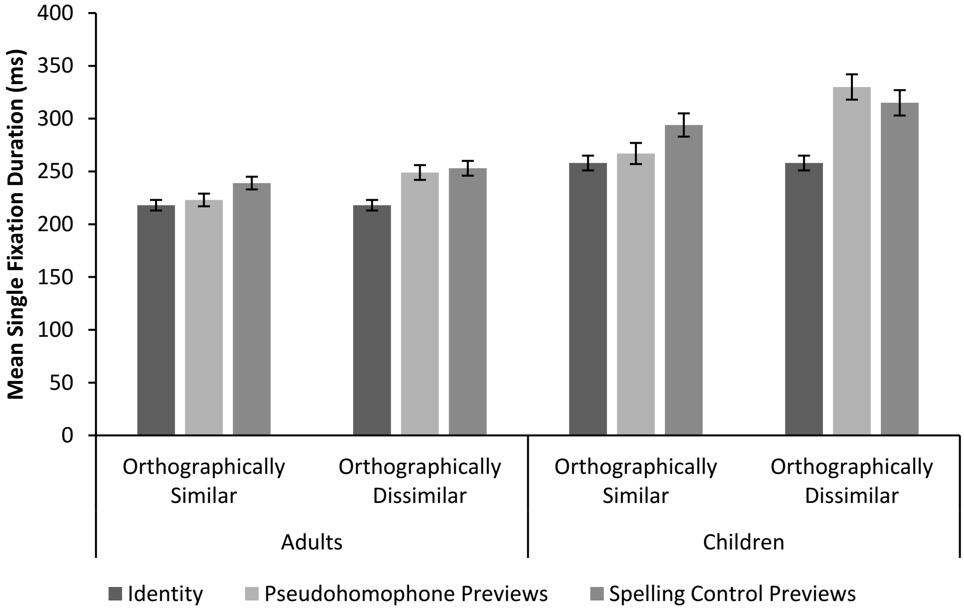
Firstly of note from this model's results, shown in Tables 4.3 and 4.4 (see also Figure 4.2), is that, in all of the measures, the children had longer reading times than the adults. Second, strikingly, in all of the measures (apart from gaze duration), the orthographically similar pseudohomophone previews gave both the adults and the children as much benefit as the correctly spelled (identity) previews. By the time, however, the readers came to move their eyes onto the next word, both the adults and the children were displaying preview benefit (faster reading times on the target word after an identity preview compared to the nonword previews). In addition, children seemed to be more affected than the adults by orthographically dissimilar pseudohomophone previews in first fixation duration; otherwise, their pre-processing was comparable.

Figure 4.2

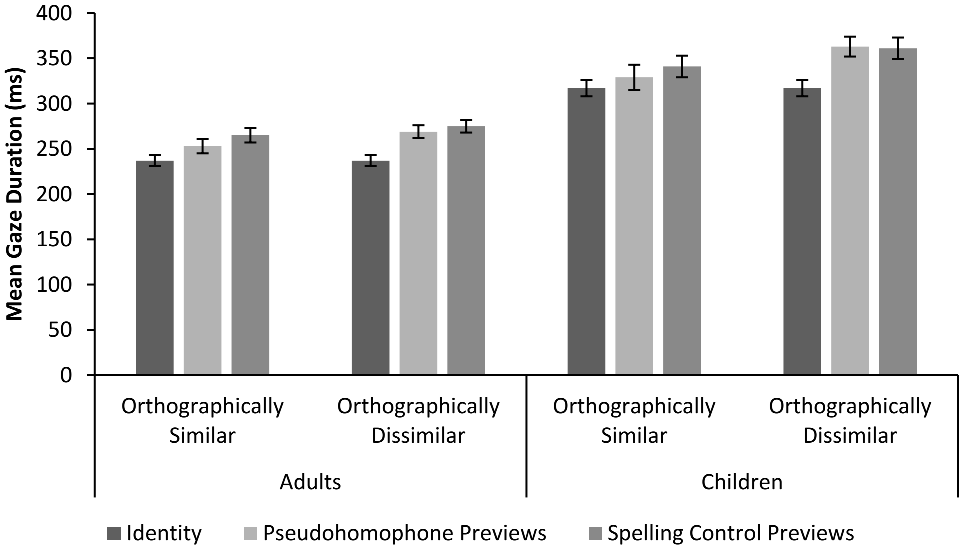
Mean First Fixation Durations (a), Single Fixation Durations (b), and Gaze Durations (c) on Identity, Pseudohomophone, and Spelling Control Previews for Both Adults and Children.



(b)



(c)



4.4.2.2 Model 2

As word length varied (stimuli word length ranged between 4-6 letters), a lme model was run with length as a factor. For all three of the dependent measures, word length had no significant effect, *t*s < +/-1.12. Formal model comparisons were also run to examine word length's role within our data. These comparisons showed, within all dependent measures, that including word length did not improve the fit of our models, *p*s > .18, consequently, we report the models without word length.

In this model, the phonological conditions were coded as: (1) pseudohomophones; and (2) spelling controls. The orthographic conditions were coded as: (1) orthographically similar; and (2) orthographically dissimilar. The *contr.sdif* function (package MASS) was used to set up the factors. Again, the reading times for adults provided the intercept for this model. The results of this model, for each of the dependent measures, are shown in Table 4.5.

Table 4.5

Output from Model 2, and contrasts, for First Fixation Duration, Single Fixation Duration, and Gaze Duration

	First fixation duration				Single fixation duration				Gaze duration			
	b	SE	*t*	*p*	b	SE	*t*	*p*	b	SE	*t*	*p*
Intercept	5.49	.02	280.78	< .001	5.53	.02	226.86	< .001	5.64	.03	219.30	< .001
Adults, Children	**.15**	**.03**	**4.79**	**< .001**	**.19**	**.03**	**5.99**	**< .001**	**.25**	**.03**	**7.67**	**< .001**
Phonological condition	.03	.02	1.64	.101	.04	.02	1.85	.065	.03	.02	1.42	.155
Orthographic condition	**.08**	**.03**	**2.70**	**.013**	**.10**	**.04**	**2.32**	**.030**	.08	.04	1.69	.105
Children × phonological condition	-.04	.03	-1.06	.290	-.02	.04	-.42	.675	-.01	.04	-.25	.801
Children × orthographic condition	.06	.03	1.66	.097	.06	.04	1.44	.150	.02	.04	.60	.552
Phonological condition × orthographic condition	-.07	.03	-1.90	.058	**-.10**	**.04**	**-2.49**	**.013**	-.05	.04	-1.25	.213
Children × phonological condition × orthographic condition	-.12	.07	-1.71	.088	-.04	.08	-.52	.601	-.01	.08	-.13	.897
Contrasts - orthographically similar previews												
Intercept	5.49	.02	245.16	< .001	5.53	.03	198.54	< .001	5.64	.03	192.11	< .001
Adults, pseudohomophone advantage	-.05	.03	-1.49	.138	**-.08**	**.04**	**-2.33**	**.020**	-.05	.04	-1.43	.152
Children, pseudohomophone advantage	**-.07**	**.04**	**-2.01**	**.045**	**-.08**	**.04**	**-2.02**	**.043**	-.05	.04	-1.19	.233
Contrasts - orthographically dissimilar previews												
Intercept	5.49	.02	244.92	< .001	5.53	.03	198.21	< .001	5.64	.03	191.85	< .001
Adults, pseudohomophone advantage	-.04	.03	-1.25	.213	-.003	.04	-.09	.929	-.009	.04	-.25	.804
Children, pseudohomophone advantage	.05	.03	1.41	.158	.02	.04	.62	.534	-.002	.04	-.04	.966

Note. The reading time data were log transformed prior to analysis, so the model estimates cannot be directly interpreted. Significant effects are

marked in bold. Following trimming, the syntax for all dependent measures for Model 2, as intercepts only models, was as follows: *depvar ~ Group * phoncond * orthcond + (1|Participant) + (1|targetno).* The contrasts were set up for all of the dependent measures within the following syntax (intercepts only models following trimming for comparing the orthographically similar pseudohomophone previews to the spelling control previews and the orthographically dissimilar pseudohomophone previews to the spelling control previews): *depvar ~ Condition2 + (1|Participant) + (1|targetno).*

As can be seen in Tables 4.3 and 4.5, and Figure 4.2, again, there were significant group differences in all measures: the children's reading times were significantly longer than those of the adults. In first fixation and single fixation durations both adults and children displayed significantly longer reading times in the orthographically dissimilar preview conditions compared to the orthographically similar preview conditions. Also, in single fixation duration, and marginally in first fixation duration, an interaction was present between the phonological conditions and the orthographic conditions. In single fixation duration both adults and children displayed a pseudohomophone advantage in preview (i.e., faster reading times in the pseudohomophone preview condition compared to the spelling control preview condition) but this was very much affected by orthographic similarity. Orthographic similarity facilitated this pre-processing of phonology, especially for the children: for the orthographically similar stimuli the pseudohomophone advantage for the adults was, on average, 16 ms and for the children it was 27 ms, whilst for the orthographically dissimilar stimuli the pseudohomophone advantage for the adults was 4 ms and for the children they actually displayed, on average, longer single fixation durations on the pseudohomophone previews than the spelling controls by 15 ms. In gaze duration, apart from the overall group differences, no significant effects or interactions were found.

Given the evident effect of orthographic similarity within some of the measures, planned contrasts were run to directly test for a pseudohomophone advantage in both the adults and the children separately within the two orthographic conditions. For each group, and dependent measure, we compared reading times on the pseudohomophone and spelling control previews in the orthographically similar and dissimilar cases. Within the orthographically similar stimuli, in both first and single fixation durations, the children were displaying a pseudohomophone advantage, and, in single fixation duration, the adults similarly benefitted from a pseudohomophone preview being present compared to a spelling control.[22]

Overall, the children were displaying a pseudohomophone advantage in their very early processing of the nonword previews that were orthographically similar to the correctly spelled identity previews (and in the adults this was significant in single fixation duration).

4.4.2.3 Bayesian Analyses

Of interest within our results was the seemingly null effect of group on condition, and the resultant null interactions. Consequently, Bayesian analyses were conducted to assess the strength of the evidence for the null and alternative hypotheses. The analyses were conducted using the BayesFactor package (Morey & Rouder, 2013), using the default scale value (.5) for the Cauchy priors on effect size and 100,000 Monte Carlo iterations. A low Bayes factor (< 1) indicates evidence for the null hypothesis, whilst a high Bayes

---

[22] Due to the multiple comparisons being undertaken within the contrasts, we also ran the contrasts (without the intercept) for each measure using the *glht* function (package multcomp) in order to adjust our *p* values accordingly (Hothorn et al., 2008). After using this correction technique, the pseudohomophone advantage displayed by the children within the orthographically similar stimuli went from being significant to marginally significant ($p = .088$ - first fixation duration; $p = .084$ - single fixation duration). We are confident, however, that even with a slight change to the significance of the children's pseudohomophone advantage within first and single fixation durations, that the children were indeed patterning in a similar way to the adults due to the Bayesian analyses conducted with regard to the null interactions present within, especially, Model 2 (see Section 4.4.2.3 in the main body of the text). The children's parafoveal pre-processing of phonology was comparable to that of the adults when the previews were orthographically similar (note that within the contrasts, in single fixation duration, even with the use of the *glht* function, the adults still displayed a significant pseudohomophone advantage within the orthographically similar stimuli, $p = .039$).

factor (> 1) provides evidence for the alternative hypothesis. Within all models items and subjects were specified as random factors. Within Model 1 we examined the null interaction between children and the orthographically similar pseudohomophone previews, comparing one model which had fixed factors of group and condition (*Group* + *Condition*) to a model which additionally had an interaction term between group and condition (*Group* + *Condition* + *Group:Condition*). The Bayes factors from the analyses were .14 for first fixation duration, .15 for single fixation duration, and .16 for gaze duration. Using the commonly cited evidence categories for Bayes factors, where a Bayes factor < .33 provides substantial evidence for a null effect, and a Bayes factor < .10 provides strong evidence, our Bayesian analyses indicate substantial evidence for the null hypothesis (i.e., the children's reading times were indeed patterning in a way consistent with the adults' reading times- displaying similar reading times between identity previews, where no display change occurred, and orthographically similar pseudohomophone previews in first fixation duration and single fixation duration).

Within Model 2, due to this model directly examining the key variables of interest, all null interactions were examined. All of the null interactions were examined by comparing a model that contained fixed factors of group and either phonological condition or orthographic condition (e.g., *Group* + *phoncond*) with a model that additionally contained an interaction term (e.g., *Group* + *phoncond* + *Group:phoncond*), and the three-way interaction was examined in a similar way (i.e., a model without the interactive terms was compared to a model with the interactive terms - *Group:phoncond:orthcond*). Regarding phonological condition, our Bayesian analyses indicated substantial evidence for the null hypothesis (.16 for first fixation duration; .12 for single fixation duration; and .10 for gaze duration), and, regarding orthographic condition, again, our analyses indicate substantial evidence for the null hypothesis (.31 for first fixation duration; .27 for single fixation duration; and .11 for gaze duration). Regarding the three-way interaction, our analyses indicate strong evidence for the null hypothesis (.01 for first and single fixation durations; and < .001 for gaze duration). Overall, the results of these Bayesian analyses suggest that the children's parafoveal pre-processing was indeed consistent with that of the adults.

4.5 Discussion

The present study investigated parafoveal pre-processing of phonology, and the potential effects of orthography on this processing, in children and adults during silent English sentence reading. The results were quite clear and, in the main, supportive of the initial predictions. Firstly, the children's reading times on the target word were significantly longer than those of the adults. This demonstrates that the children experienced greater processing difficulty, that is, a slower rate of lexical processing, during reading than the adults, consistent with past research (e.g., Milledge et al., 2020). This has also been shown in simulations of the E-Z reader reflecting differences between adults' and children's eye movement behaviour during reading (Reichle et al., 2013). Secondly, the children, like the adults, were sensitive to manipulations of orthography in preview, consistent with past research (e.g., Johnson et al., 2018; Milledge et al., 2020; Pagán et al., 2016; Pollatsek et al., 1992). Thirdly, the adults did display evidence of a pseudohomophone advantage within the orthographically similar stimuli (albeit only significant in one measure, as discussed later). Critically, though, this study provides evidence that phonology also played a role in preview for 8- to 9-year-old readers of English; intact phonological codes present in preview facilitated lexical processing of word *n*+1. Interestingly, and in line with our predictions for the adults, there was evidence of an interactive effect between orthography and phonology and, with novelty, this was also found within the children: reading times were generally faster the greater the degree of orthographic similarity between the nonword previews and the correct target word, and, importantly, this effect was augmented for nonwords that maintained the target word's phonology in preview.

Evidently, phonology had a significant early effect on both the adults' and, especially, the children's pre-lexical processing. Both the adults and the children gained as much benefit from the orthographically similar pseudohomophone previews as the identity previews (where no display change occurred) very early in their processing. The orthographically similar previews with, critically, intact phonology facilitated readers' lexical identification as much as the correctly spelled target word being present in preview. For the orthographically similar previews, the letter substitutions occurred towards the end of the words, and it has been shown that children are sensitive to letter substitutions across the whole-word form in the parafovea (up to six letters), with letter substitutions near the ends of target words increasing children's reading times during direct fixation, compared to an identity preview (Milledge et al., 2020). Thus, even though the children would have been able to extract orthographic information from those letter positions (i.e., a cost would have been expected due to the manipulation of orthography- letter substitutions- in preview), they gained the same processing benefit as if they had the full, unmanipulated target word (i.e., the identity) in preview so long as those substitutions preserved the word's phonology.

Moreover, in two measures of very early processing, the children displayed a pseudohomophone advantage for the orthographically similar previews, and in one measure the adults showed this advantage too. Within the orthographically similar stimuli, when the phonology of the target word was maintained in preview, this facilitated children's (and adults') processing significantly, and to a significantly greater degree than was the case for a spelling control preview. The children clearly benefitted from correct phonological information being present in preview, and this facilitated lexical identification of the target, although orthographic similarity was evidently also playing a role. These results extend, and complement, existing research findings regarding children's, and adults', phonological and orthographic processing during silent sentence reading. Whilst such research has suggested that phonological processing is pre-lexical in child readers of English (Blythe et al., 2015; Jared et al., 2016), this study provides the first evidence of child readers extracting phonological information from an upcoming word through parafoveal pre-processing.

Importantly, as briefly mentioned previously, we found evidence of an interactive effect between phonology and orthography in preview. Whilst this replicates a known effect in adults (Pollatsek et al., 1992), this is novel for children. Interestingly, this indicates that the children were engaging in quite sophisticated parafoveal pre-processing, comparable to that of the skilled adult readers. The adult and beginner 8-year-old readers showed remarkable similarities in their parafoveal pre-processing: processing the same information from word n+1 that subsequently facilitated their lexical identification of that word during direct fixation. Indeed, the interaction found between phonology and orthography demonstrates that, for both adult and child readers, the ability to undertake phonological processing of word n+1 was modulated by orthographic similarity during early processing. The greater the overlap between the orthographic code(s) with the phonological code(s) of word n+1, potentially driven by the first letter, the greater the facilitation to lexical identification processes (consistent with Grainger & Ziegler, 2011). Parafoveal pre-processing would appear to be a skill that has, therefore, largely developed even in 8-year-old children to be similar to that of skilled adult readers, qualitatively; however, some quantitative differences remain, with children's reading and lexical identification processes being slower and less efficient than those of the adults. Presumably this continues to change developmentally, as beginner readers progress to be skilled readers and develop higher quality lexical representations (e.g., Perfetti, 2007).

The present experiment's findings are consistent with Grainger and Ziegler's (2011) model of orthographic processing, though we note again that the model is based on identification of directly fixated words presented in isolation and we are making inferences about how this might extend to parafoveal pre-processing during sentence reading. It is implicit within this model that there is a developmental change in lexical identification strategy from overt, effortful phonological decoding to the use of whole-word orthographic encoding (coarse-grained and fine-grained). Importantly, within this orthographic encoding, a mechanism is retained that allows phonological representations to be activated pre-lexically within the fine-grained route (phonological recoding; i.e., phonological information to be processed from word *n*+1). Consequently, both the adults (as would be expected) and the children appeared to be using the fine-grained route of processing: they were both able to rapidly, pre-lexically, extract phonological information from word *n*+1. As posited by this theory, though, orthography was also having an effect. Within the fine-grained route, as stated previously, there is little flexibility regarding orthographic encoding, and, as such, there is little tolerance for word misspellings. A pseudohomophone advantage was found, therefore, within the orthographically similar previews but not within the orthographically dissimilar previews: whilst word misspellings were present within both types of preview, within the orthographically similar previews the misspellings were clearly better tolerated, due to their lesser disruption to the orthographic processing the participants had to undertake (as only one letter was substituted in preview, never the first letter), in comparison to the orthographically dissimilar previews and their greater disruption to participants' orthographic processing (as at least two letters were substituted in preview, for half of the previews this involved at least the first letter). To be clear, we consider such effects might be a reflection of the importance of the first letter to both adult and child readers' parafoveal pre-processing, with its intact orthography facilitating the extraction of phonology from word *n*+1.

Indeed, the present study suggests that the adults and, especially the children, might have been displaying a first-letter bias in their parafoveal pre-processing, as shown previously by Milledge et al. (2020). Milledge et al. found, broadly consistent with past research (e.g., White et al., 2008), that the first letter played an important role in adults' pre-processing (the adults displayed longer reading times when the first letter was substituted in preview relative to the identity condition, where all letters were maintained in preview). In addition, Milledge et al. found that children, very early in their lexical processing (first fixation duration), displayed a first-letter bias: longer reading times after previews where the first letter was substituted relative to previews where the end letter was substituted. Within the present experiment, similarly, in first fixation duration the children's reading times were significantly more affected than the adults' by the orthographic similarity of the nonword previews to the correct target word, with this effect appearing to be especially apparent with regard to the orthographically dissimilar pseudohomophone previews. Of note, to reiterate, is that within the orthographically dissimilar previews at least one of the letters being substituted would be the first and/or second letter (e.g., *kween* as a preview for *queen*), in contrast to the orthographically similar pseudohomophone previews where only one letter would be substituted, and it was never the first or second letter (e.g., *cheeze* as a preview for *cheese*). These substitutions of the first and/or second letters involved in the orthographically dissimilar pseudohomophone previews seems to have come at a particular cost to the children suggesting that, despite the correct phonological codes being present in preview, the children were not able to benefit from this information due to the letter substitutions occurring near the beginning of target words in preview. With half of the previews within the orthographically dissimilar pseudohomophone condition involving at least the first letter of the correct target word being substituted in preview, the increased disruption the children experienced to their reading times, very early in their processing (similar to

Milledge et al., 2020), could potentially be attributable to this orthographic manipulation of the first letter/s in preview. Overall, given the interactive relationship found between orthography and phonology in both the adults and the children, the results suggest that the adults were similarly affected to the children: the greater the overlap between orthography and phonology, the more able readers were to benefit from phonological information in preview, with the first letter potentially playing a pivotal role in this overlap.[23]

Like the children, the adults did experience benefits from correct phonological information being present in preview, broadly consistent with past research (e.g., Pollatsek et al., 1992). It is worth noting, however, that these benefits with regards to displaying a pseudohomophone advantage were mainly only present in numerical trends within our data. The pseudohomophone advantage was small in the adults and was not consistent across the early measures of processing (i.e., only significant in single fixation duration), unlike in the children. It is of note, though, that when looking at proportional increases within the orthographically similar stimuli, the mean costs between the pseudohomophone and spelling control previews were not that different between the adults and the children: first fixation duration, 6% for the adults, 9% for the children; single fixation duration, 7% for the adults, 10% for the children. This suggests that processing within the adults and the children was largely comparable, given the similar proportional increases. We consider that the small effect found within the adults in our formal analyses could be due to the stimuli used. Similar to Milledge et al. (2020) and Tiffin-Richards and Schroeder (2015a),

---

[23] We sought to investigate this by running contrasts comparing the orthographically dissimilar pseudohomophone previews where the first letter was manipulated in preview (e.g., *kween*) to the orthographically dissimilar pseudohomophone previews where the first letter was not manipulated in preview (e.g., *hunni*). No significant effects were found, probably due to the small number of stimuli involved. The means (available in Appendix B; B.2; Table B.2.3) are supportive, though, of the cost of the first letter being substituted in preview, especially for the children.

the stimuli used were designed to be suitable for the given age-group of child readers, not skilled adult readers. Consequently, the sentences would have been very easy for the skilled adults to read. This ease of processing may have resulted in the adults allocating more attention to processing of upcoming words within a sentence than they would have been able to do with more demanding, age-appropriate, sentences (e.g., Henderson & Ferreira, 1990; Rayner, 1986; though see also Zhang et al., 2019). As a result, smaller differences would have been found in reading times between the preview conditions. Thus, whilst the adults did display the predicted pattern of results numerically, significant effects were less likely to be found, given their greater ability (in comparison to the beginner child readers) to allocate more attentional resources towards pre-processing word *n*+1. Interestingly though, even in studies using stimuli designed for adult readers of English, the effect of phonology (pseudohomophones/homophones) in preview is typically small, about 4 ms in gaze duration, with little evidence of an effect of phonology in first fixation duration (Vasilev et al., 2019). The fact that we found a pseudohomophone advantage in the adult readers in an earlier measure of processing than gaze duration- single fixation duration- supports the notion that the stimuli, and the ease of the adults' processing, were potentially behind this effect. The adult readers seemed to be gaining an early advantage from phonology in preview (from orthographically similar previews), that is, before their eyes left a target word for the first time, but, by the time their eyes had moved onto the next word, they were no longer significantly displaying this effect.

Consistent with teenage readers (Blythe et al., 2018, 2020), the typically developing 8- to 9-year-old children were undertaking covert, rapid phonological recoding during their silent sentence reading. Although this has been suggested by past research investigating foveal processing (Blythe et al., 2015; Jared et al., 2016), this is the first experiment that has provided direct evidence of this through examining pre-lexical, parafoveal (pre-)processing. Clearly, typically developing 8- to 9-year-old beginner readers of English have made the transition from phonological decoding to recoding: they have moved beyond the slow, effortful sounding out of letters to identify a word to the rapid, pre-lexical processing of phonology, as demonstrated by their ability to lexically identify an upcoming word being facilitated by correct phonological information being present in preview (i.e., demonstrating pre-lexical processing). Whilst phonological decoding is a phase included in most theories of learning to read (e.g., Ehri, 1995, 1998, 1999, 2005, 2007; Marsh et al., 1981; Mason, 1980), what is unclear is exactly how, and when, beginner child readers make this transition from phonological decoding to recoding. Although the present experiment does not shed light on how exactly this transition occurs, the results do suggest that this transition has occurred at least by the time typically developing readers of English are 8-years-old. Future research could examine this issue. Given the ability to extract phonological information from the parafovea is dependent on the development of phonological recoding, due to the pre-lexical nature of this processing, it would be expected that younger child readers who have not made this transition would not show the same preview effects (i.e., they would not display a pseudohomophone advantage). In relation to Grainger and Ziegler's (2011) model, typical child readers of English, as young as 8-years-old, appear to have developed phonological processing, within their fine-grained route, that is comparable in efficiency to that of skilled adult readers: they can undertake rapid, covert, pre-lexical processing of phonology (phonological recoding). Younger child readers, however, who are reliant on the lexical, foveal strategy of phonological decoding should not be able to extract phonological information from an upcoming word, that is, display pre-lexical, parafoveal processing of phonology.

In sum, the current experiment provides novel evidence of 8- to 9-year-old beginner readers of English parafoveally pre-processing phonology, in a broadly similar way to skilled adult readers. Both groups displayed evidence of undertaking covert, pre-lexical phonological recoding. Of note also, though, is the key role orthography appears to play in facilitating this pre-processing of phonology.

Chapter 5 The Importance of the First Letter in Children's Parafoveal Pre-processing in English: Is it Phonologically or Orthographically Driven?

The data that supports the findings of this study, and the code used for the analyses, are available from: https://osf.io/p8mh7/?view_only=5fe2585f5b244dbc8ad621358766a334.

5.1 Abstract

For both adult and child readers of English, the first letter of a word plays an important role in lexical identification. Using the boundary paradigm during silent sentence reading, we examined whether the first-letter bias in parafoveal pre-processing is phonologically or orthographically driven, and whether this differs between skilled adult and beginner child readers. Participants read sentences which contained either: a correctly spelled word in preview (identity; e.g., *circus*); a preview letter string which maintained the phonology, but manipulated the orthography of the first letter (P+ O- preview; e.g., *sircus*); or a preview letter string which manipulated both the phonology and the orthography of the first letter (P- O- preview; e.g., *wircus*). There was a cost associated with manipulating the first letter of the target words in preview, for both adults and children. Critically, during first-pass reading, both adult and child readers displayed similar reading times between P+ O- and P- O- previews. This shows that the first-letter bias is driven by orthographic encoding, and that the first letter's orthographic code in preview is crucial for efficient, early, processing of phonology.

5.2 Introduction

A word's orthography (its printed form) and phonology (its associated speech sounds) are inherently linked within alphabetic languages, though it is of note that this does vary based on orthographic depth and how consistent grapheme-to-phoneme correspondences (GPCs) are within a language (e.g., Katz & Frost, 1992). Nevertheless, within English, orthographic information visually represents the phonological codes of a word in order for word (lexical) identification to occur during reading. Past research has shown that the first letter of a word appears to play a vital role within both adults', and especially, children's lexical identification processes in English, facilitating lexical identification of an upcoming word (Milledge et al., 2020, 2021a). It is unknown, however, exactly what drives this first-letter bias. In the present study, we examined the first-letter bias in 8- to 9-year-old readers of English, seeking to determine whether this bias might be phonologically or orthographically driven.

It is well-documented that, during silent reading, readers begin to process the upcoming word (*n*+1) in the sentence whilst still fixating the current word (*n*) (see Rayner, 1998, 2009 for reviews). This is referred to as parafoveal pre-processing, and leads to faster reading times for word *n*+1 when it is directly fixated due to the processing that has already occurred in relation to that word. It is typically studied using the boundary paradigm (Rayner, 1975). In this paradigm, an invisible boundary is placed immediately before a target word. Prior to the readers' eyes crossing this boundary, a preview letter string is present in place of the correct target word. When the readers' eyes cross this boundary, the preview letter string changes to the correct target word. Faster reading times on the target word following a correct preview (i.e., an identity condition, where the preview letter string is identical to the correct target word) compared to other preview letter strings (experimental conditions where the preview has been manipulated to be different in some manner) is known as preview benefit (see Schotter et al., 2012 for a review). Through systematic variation of the preview letter string in relation to the target word, researchers are able to determine the type of information readers extract and use from word *n*+1.

Past research using the boundary paradigm has shown that a word's external letters (both beginning and end) are particularly important to skilled adult readers' parafoveal pre-processing and subsequent lexical identification. Further, the first letter of a word plays a more privileged role than the end letter, during both parafoveal pre-processing and subsequent direct fixation (e.g., Briihl & Inhoff, 1995; Inhoff, 1987, 1989a,b; Johnson & Eisler, 2012; Johnson et al., 2007; Rayner et al., 1980; White et al., 2008). For example, White et al. (Experiment 1; 2008) found that reading times were slower when a word was present with external letter transpositions (e.g., *problme*, *rpoblem*) compared to internal letter transpositions (e.g., *porblem*, *probelm*). Within the external letter transposition conditions, however, reading times were slower when the transpositions occurred at the beginning relative to the end of a word (e.g., *rpoblem* vs. *problme*). The same pattern of effects was also observed when parafoveal pre-processing of the target word was prevented, through preview of the word to the right of fixation being unavailable (Experiment 2). This suggests that the first letter of a word plays a critical role in both parafoveal pre-processing and foveal lexical identification for skilled adult readers.

Similar effects have also been found within beginner child readers. Milledge et al. (2020) found that in children, like adults, the manipulation of external letters in preview was more detrimental to their lexical processing than the manipulation of internal letters in preview (e.g., *romter*, *sislun* vs. *somler*, *simlur* as previews for *sister*). Moreover, both adults and children experienced a clear cost when parafoveal pre-processing of the first letter was denied. This first-letter bias occurred earlier during lexical processing for children than was the case for effects of other letter manipulations. For the majority of effects reported, the time course was delayed in children compared to adults (e.g., not present in first fixation duration but present in gaze duration and total reading time). In contrast, when the first letter was substituted in preview, both the adults' and the children's parafoveal pre-processing was immediately, and similarly, disrupted. Evidently, the beginning letter of a word plays an important role in facilitating both children's and adults' lexical identification of word *n*+1, given the cost to their reading times when this letter is disrupted in preview.

This first-letter bias is a robust finding within the literature, but it is unclear as to what causes it. Within skilled adult readers, the possibility that this effect occurs due to fundamental constraints of the visual system, like visual acuity and lateral masking, can be rejected. For example, Johnson and Eisler (2012) found that when lateral masking was equated for all letters of a word through the replacement of inter-word spaces with #s (e.g., The#boy#could#not#solve#the#problem#so#he#asked#for#help.), word initial letter transpositions still caused more disruption to reading than word final transpositions, whilst the end letter transpositions were no more disruptive than internal letter transpositions (Experiments 1 and 2). Furthermore, manipulations of the first letter of a word remain particularly disruptive even when participants are required to read sentences backwards, from right to left (e.g., *.help for asked he so problem the solve not could boy The*) (Experiment 4; Johnson & Eisler, 2012). Within such sentences, during fixation on word *n* (e.g., *the*), the first letter of word *n*+1 (e.g., the *p* in *problem*) falls furthest away from the point of fixation. The first letter of the word being pre-processed will, therefore, be perceived with the lowest visual acuity within that word, whilst the final letter (e.g., the *m* in *problem*) will be perceived with the highest visual acuity, as it falls closest to the point of fixation on word *n*. Even under such conditions, manipulations of the first letter in preview (e.g., *rpoblem*) were more disruptive to reading than manipulations of the last letter in preview (e.g., *problme*). Consequently, visual factors, like lateral masking and the proximity of the first letter to the point of fixation (visual acuity), do not seem to play a causal role in the importance of the first letter to lexical identification. This suggests that the first-letter bias may be driven by cognitive processing associated with lexical identification. However, this leads to the question of whether the parafoveal pre-processing that operates over the word initial letters is associated with the extraction of orthographic or phonological information.

First, it could be occurring as part of orthographic encoding, given the effect orthographic manipulations of the first letter have on both adults' and children's ability to lexically identify a word (e.g., Milledge et al., 2020). Alternatively, the effect could be caused by the reader's generation of a phonological code, which necessarily requires left-to-right processing of the letters within a word. Skilled adult readers pre-process phonological codes from word *n*+1 as part of lexical identification in silent sentence reading, as shown in a number of studies that have used the boundary paradigm. For example, adults display faster reading times after a homophone preview (e.g., *beech* as a preview for *beach*) compared to a spelling control preview (e.g., *bench* as a preview for *beach*; Chace et al., 2005; Pollatsek et al., 1992). Similarly, it has been found that beginner child readers also extract phonological information from word *n*+1; for example, through displaying faster reading times on a target word after a pseudohomophone preview (e.g., *cheeze* as a preview for *cheese*) compared to a spelling control preview (e.g., *cheene* as a preview for *cheese*) (Milledge et al., 2021a). It is possible, therefore, that the first-letter bias could be phonologically driven in adult and child readers.

The first-letter bias has been accounted for by various models of word recognition, though we note that these models typically relate to isolated word recognition under direct fixation, not during parafoveal pre-processing (e.g., Davis, 2010a,b; Grainger & Ziegler, 2011; Perry et al., 2007; Whitney, 2001). For example, the SERIOL (Whitney, 2001) and Spatial Coding (Davis, 2010b) models of visual word recognition both account for this importance of the first letter: within the SERIOL model, given left-to-right scanning of a word, letters in the first position receive the most activation; within the Spatial Coding model, dynamic end-letter marking is used, such that the first and final letters of a word are weighted more heavily than other constituent letters of a word. Given the nature of these models, and how they relate to letter position encoding, not only do they predict that the first letter plays a vital role in lexical identification, but that this role is, first and foremost, orthographically driven.

Despite the models' focus upon isolated word recognition, they do have the potential to provide insight into lexical identification processes, regarding word *n*+1, during natural sentence reading (e.g., Pagán et al., 2016). Consequently, insight can be gained from such models into why the first-letter bias occurs and what may drive this effect. For example, Grainger and Ziegler's (2011) model of word recognition proposes that there are two processing routes through which lexical identification can be achieved: coarse-grained and fine-grained. The coarse-grained processing route gives a reader access to semantics (meaning) from a word's orthographic form. The fine-grained route, in contrast, provides a reader access to semantics through the processing and mapping of commonly occurring letter patterns onto their corresponding phonological representations. Whilst the former route allows some flexibility with regard to orthographic encoding, the latter route allows little flexibility with regard to orthographic encoding (i.e., reduced tolerance of word misspellings). Specifically, it would appear that the first letter's correct orthographic code plays a particularly important role within the orthographic encoding readers undertake in the fine-grained route to lexically identify a word, facilitating efficient processing of phonology (Milledge et al., 2021a). To be clear, given the supposition that phonological lexical representations are accessed via encoding and recognition of corresponding orthographic form/s (Perry et al., 2007), the first letter of a word may be crucial for readers to translate an orthographic code into a phonological code. This would suggest that the first-letter bias is potentially, primarily, orthographically driven, rather than phonologically driven.

In the present study, we examined whether the first-letter bias in parafoveal pre-processing is orthographically or phonologically driven by manipulating the features of the first letter of target words in parafoveal preview. We compared the effects of these manipulations, and their time course, for beginning and skilled adult readers. Previews were either: the correct target word (identity; e.g., *circle*); a letter string with the first letter substituted such that the phonological code of the first letter was maintained (P+ O-; e.g., *sircle*), or a letter string with the first letter substituted such that both the phonological and orthographic codes were disrupted (P- O-; e.g., *nircle*).

First, we predicted that both children and adults would show a cost to their processing when the first letter was substituted in preview, compared to the identity condition (Milledge et al., 2020, 2021a).[24] Second, we predicted that a comparison of the two substitution conditions would indicate the cause of the first-letter bias. Specifically, if the effect is phonologically based, then we would expect shorter reading times after a preview where the phonological code of the first letter was preserved (e.g., faster reading times on a target word after a P+ O- preview compared to a P- O- preview). Alternatively, if the first-letter bias is orthographically driven then we would expect both substitution conditions to have similar reading times. We also predicted that, in addition to overall group differences (i.e., the children displaying longer reading times than the adults; e.g., Blythe & Joseph, 2011), differences in the time course of effects were likely to be found

---

[24] Although adult readers do not necessarily show a traditional first-letter bias in their pre-processing where stimuli designed for children are used, they do still seem to use the first letter as an important cue within their pre-processing (Milledge et al., 2020).

between the adults and the children; in particular, differences- delays- within the children's parafoveal pre-processing of orthography (Milledge et al., 2020).

5.3 Method

5.3.1 Participants

Forty-two adults (*M* = 22.17, *SD* = 3.15) and 42 8- to 9-year-old (*M* = 8.43, *SD* = .50) children from a local junior school participated in the eye-tracking experiment (see Table 5.1 for a summary of group characteristics). All were native speakers of English, had normal or corrected to normal vision, and no known reading difficulties, as confirmed by the Wechsler Individual Achievement Test II UK (WIAT-II UK; Wechsler, 2005) reading subtests. All participants' composite standardised scores were within the expected range (adults' score range: 92-134; children's score range: 95-142; see also Table 5.1). Ethical approval was provided by the University of Southampton Psychology Ethics Committee (submission ID: 52927.A1).

Our sample size is broadly comparable to other studies on children's parafoveal pre-processing (e.g., Marx et al., 2015; Milledge et al., 2020, 2021a; Pagán et al., 2016; Tiffin-Richards & Schroeder, 2015a). Regarding the number of stimuli, we exhausted all possibilities and note that the number of stimuli per condition per participant, as outlined in the following subsection, is greater than most previous experiments: 17 in the present experiment vs. e.g., eight (Pagán et al., 2016); 10 (Marx et al., 2015); 10- orthographic manipulation- and 14- phonological manipulation (Tiffin-Richards & Schroeder, 2015a).

Table 5.1

Summary of Group Characteristics

		Mean	StDev	*t*	*df*	*p*
Test age (years)	Adults	22.17	3.15			
	Children	8.43	.50	27.88	82	< .001
WIAT word reading	Adults	111.60	4.57			
	Children	111.86	10.37	-.15	82	.881
WIAT pseudoword decoding	Adults	107.14	8.86			
	Children	107.17	8.00	-.01	82	.990
WIAT comprehension	Adults	113.81	5.63			
	Children	115.07	7.70	-.86	82	.394
WIAT composite standardised scores	Adults	115.10	9.06			
	Children	112.90	11.44	.97	82	.334

Note. The three right-hand columns give the results of independent samples *t*-tests comparing the adults to the children. All WIAT scores are standardised.

5.3.2 Materials and Design

We selected 24 potential 5-7 letter target words, which were either nouns or adjectives. These target words were selected on the basis that the first letter of each of the words could be substituted with an orthographically similar letter (e.g., a descender replaced with a descender), in order to create a preview letter string that would maintain the phonology of the target word (e.g., a pseudohomophone). This was done due to the interactive relationship between orthography and phonology (Milledge et al., 2021a); specifically, the orthographic dissimilarity of the first letter in preview (e.g., *c* substituted with *k*; *kley* as a preview for *clay*) could play a role in further disrupting readers' ability to extract phonological information from word *n*+1, given how orthography has been found to be pre-processed by children (e.g., Johnson et al., 2018; Pagán et al., 2016). Consequently, and given the constraints within the English language, all target words either began with a *c* that could be substituted with a *s* in preview to give the first letter its correct phonological code (e.g., *sircle* as a preview for *circle*) or a *g* that could be substituted with a *j* in preview (e.g., *jiraffe* as a preview for *giraffe*).

For each of the 24 target words, four potential sentence frames were created. All materials were pre-screened for both the difficulty of the sentences and whether the given target words were known and recognised by the target age group. Forty-five 8- to 9-year-old children (all of whom were native speakers of English with no known reading difficulties, and none of whom took part in the eye-tracking experiment) rated the sentences on a scale of 1 (easy to understand) to 7 (difficult to understand). The children were also asked to underline any words in the sentences that they did not know or recognise. The target words and sentence frames were selected to ensure that they were easy for our target age group to understand (had a mean rating under 2.00) and on the basis of the target words being known by all of the children. As a result of this pre-screening, seven target words and their associated sentence frames were dropped. This left a final stimulus set of 17 target words (the linguistic properties of these words are shown in Table 5.2). For each of these target words, three sentence frames were chosen for the eye-tracking experiment; the sentence rated as most difficult, on average, out of the four potential sentence frames was dropped. Consequently, the final stimulus set consisted of 51 experimental sentences (see Appendix A; A.2).

Table 5.2

Linguistic Properties of the Target Words and Sentence Frames

	Target words
Orthographic neighbours (N-Watch; Davis, 2005)	≤ 2
Age of Acquisition (Kuperman et al., 2012)	*M* = 6.28 years *SD* = 1.56
Child frequency counts (Children's Printed Word Database; Masterson et al., 2003)	Range = 3-430 per million *M* = 55 *SD* = 103
Adult frequency counts (English Lexicon Project Database; HAL corpus, Balota et al., 2007)	Range = 379-148,204 per million *M* = 26,531 *SD* = 38810
Understandability (1 *easy* to 7 *difficult*)	Range = 1-1.53 *M* = 1.17

Note. The adult frequency counts refer to 16 of the target words (*gerbil* was not available in the database).

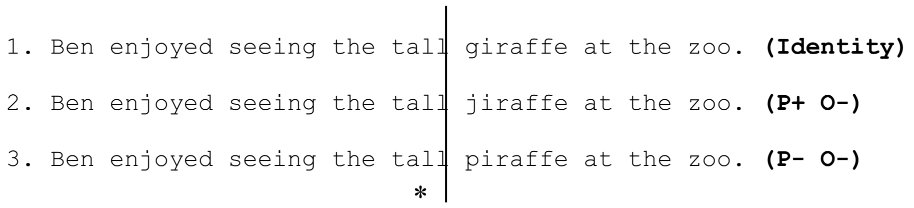
The gaze-contingent boundary paradigm (Rayner, 1975) was used. In the present experiment, three parafoveal preview conditions were generated for each target word. There was an identity (control) condition, where the preview letter string was identical to the correct target word (e.g., *giraffe* - *giraffe*), and two experimental conditions, which involved the substitution of the first letter of each of the target words in preview: P+ O- previews (where the correct phonological code of the first letter was maintained in preview and orthography was manipulated; e.g., *jiraffe* - *giraffe*) and P- O- previews (where both the phonological and orthographic codes of the first letter were manipulated in preview;

e.g., *piraffe* - *giraffe*). All nonwords were orthographically legal and pronounceable. The P+ O- and P- O- previews were matched on bigram and trigram frequency, as well as orthographic neighbourhood size (the number of real words that could be formed by making a single, position-specific letter substitution), *t*s < .59 (N-Watch; Davis, 2005).

Every participant read all of the 51 experimental sentences, contributing data to all three preview conditions, and 17 filler sentences were also included. Consequently, there were 17 stimuli per condition per participant. As every participant saw each target word three times, and was provided with three different previews, within the 51 experimental sentences the preview condition presentation order was carefully controlled: six files were created accounting for each possible combination of preview presentation (i.e., whether a given participant had an identity preview of a given target word on first, second, or third presentation, or a P+ O- preview, or a P- O- preview). The sentences occupied one line on the screen (maximum = 55 characters; *M* = 50 characters; e.g., *Ben enjoyed seeing the tall giraffe at the zoo.*). An example stimulus in each condition is shown in Figure 5.1.

Figure 5.1

Example of an experimental sentence showing the three parafoveal preview conditions, with the invisible boundary shown (though shown as a visible line here it was not visible to participants). The asterisk refers to a fixation, the target word is shown in bold, and the condition is shown in the brackets.



1. Ben enjoyed seeing the tall giraffe at the zoo. **(Identity)**

2. Ben enjoyed seeing the tall jiraffe at the zoo. **(P+ O-)**

3. Ben enjoyed seeing the tall piraffe at the zoo. **(P- O-)**

5.3.3 Apparatus and Procedure

An EyeLink 1000 eye-tracker recorded eye movements of the right eye (SR Research). Forehead-and-chin rests were utilised to minimise head movements. A three-point calibration and validation procedure was carried out. The procedure would be repeated if the mean validation error, or the error for any of the individual points, was greater than .20°. A single sentence was presented to participants at a time in black, Courier New, 14-point font on the grey background of a 21in. CRT monitor, which had a refresh rate of 120 Hz. The viewing distance was 60 cm; one character subtended .34° of visual angle. Participants were instructed to read silently and for comprehension. In order to familiarise participants with the procedure, they were presented with four practice trials at the beginning of the experiment (with two comprehension questions). After finishing reading a sentence, participants would press a response key, and one-third of the sentences were replaced by a yes/no comprehension question to which the participants would have to respond. After the eye-tracking, participants were asked if they had noticed anything strange about the sentences they had been reading, as detecting display changes can affect fixation times (e.g., White et al., 2005). Six adult participants' data was excluded from the analyses on this basis and were replaced with adult datasets where no display changes were detected. Participants then completed the three reading subtests of the WIAT-II UK (Wechsler, 2005). The whole experiment lasted about 50 minutes per participant.

5.4 Results

All participants scored at least 76% on the comprehension questions (adults: *M* = 98.32%, *SD* = 2.99%; children: *M* = 93.84%, *SD* = 7.11%). The data were trimmed using the clean function in DataViewer (SR Research). Fixations shorter than 80 ms, and which were located within one character space of a neighbouring fixation, were merged into the neighbouring fixation. Remaining fixations that were shorter than 80 ms or longer than

1,200 ms were deleted. In total 1,370 fixations were merged or deleted (2.25% of the dataset; 637 adult fixations and 733 child fixations), resulting in a final dataset of 59,509 fixations.

All data were analysed using linear mixed effects (lme) models, using the *lmer* function from the lme4 package (Bates et al., 2015) within the R environment for Statistical Computing (R Core Team, 2020). Participants and items were entered as crossed random effects. For each model, full random structures were initially specified for items and participants, to avoid being anti-conservative (Barr et al., 2013). Failure of the models to converge for each dependent measure led to the models' structures being trimmed until they would converge. Data (for both global and local analyses) were log transformed before analysis to reduce skew.[25]

5.4.1 Global Measures

Firstly, we examined global measures of participants' eye movement behaviour (eye movements across entire sentences). As can be seen in Table 5.3, the children displayed significantly longer fixation durations ($b = .10$, $SE = .02$, $t = 4.61$, $p < .001$), longer total sentence reading times ($b = .53$, $SE = .07$, $t = 7.31$, $p < .001$), and made more fixations ($b = .39$, $SE = .05$, $z = 7.20$, $p < .001$) than the adults, consistent with previous research (e.g., Blythe et al., 2011; Blythe & Joseph, 2011; Joseph et al., 2009; Tiffin-Richards & Schroeder, 2015a,b).[26]

---

[25] Note that, within the global analyses, due to the nature of the fixation count data it was not log transformed and was analysed using a generalized linear mixed model, in order to use the Poisson distribution.

[26] Following trimming, the syntax for fixation count was: *Fix_count ~ Group + (1|Participant) + (1 + Group|SentenceNo)*, the syntax for total sentence reading time was: *Total_sentence_reading ~ Group + (1|Participant) + (1 + Group|SentenceNo)*, and the syntax for fixation duration, as an intercepts only model, was: *Fix_duration ~ Group + (1|Participant) + (1|SentenceNo)*.

Table 5.3

Mean and Standard Deviation (in parentheses) Values for Measures across Entire Sentences

Measure	Adults	Children
Fixation duration (ms)	239 (116)	273 (150)
Fixation count	11 (4)	17 (6)
Total sentence reading time (ms)	2624 (1141)	4635 (2413)

5.4.2 Local Measures

Subsequently, we analysed reading time data on the target word in each sentence. Before analysing the local dependent measures, the data were further cleaned: trials were excluded from the analyses if the boundary change occurred early during a fixation on the pretarget word and if the boundary change was late- not completed until more than 15 ms after fixation onset on the target word (224 adult trials- 10.46% of the adult trials, and 202 child trials- 9.43% of the child trials).[27]

The key dependent measures were: first fixation duration (the duration of the first fixation on a word, irrespective of how many fixations the word received), single fixation

---

[27] A late boundary change was also operationalised as 10 ms, in order to ensure that the pattern of data remained unchanged between the two reports (i.e., 10 ms vs. 15 ms report). The pattern of data was highly consistent across the two sets of analyses that were conducted using these reports, for all measures, so the 15 ms criterion of a late boundary change was used as it allowed the retention of more data (3,858 data points compared to 3,672). After the boundary change cleaning, regarding the total number of items recorded for each participant, within the adults the lowest total number of items recorded was 35 (M = 45.67, total range: 35-51; identity M = 15.05, range: 9-17; P+ O- M = 15.69, range: 12-17; P- O- M = 14.93, range: 11-17) and within the children this was also 35 (M = 46.19, total range: 35-51; identity M = 15.57, range: 11-17; P+ O- M = 15.05, range: 11-17; P- O- M = 15.57, range: 11-17).

duration (the duration of the first fixation on the word when it received only one first-pass fixation), gaze duration (the sum of all first-pass fixations on a word before the eyes move from that word), selective regression path duration (the sum of all fixations made from the moment the eyes land on a target word until the first fixation to the right of the target word, not including time spent rereading preceding text), and total reading time (the sum of all fixations made on a target word); see Table 5.4.[28]

---

[28] Across all trials, the probability of the children making a single fixation was .56 and the probability of the adults making a single fixation was .77. Single fixation probabilities for the adults and the children by condition are available in Appendix B; B.3 (Table B.3.1). Within Appendix B; B.3, skipping rates are also provided in Table B.3.2. The only significant finding from the generalized linear mixed models conducted for this measure was that, within Model 1 (intercepts only model), the adults were significantly less likely to skip a P- O- preview than an identity preview (*p* = .011), and the lack of significant interaction term suggests that the children's skipping behaviour was consistent with that of the adults (*p* = .262).

Table 5.4

Mean and Standard Deviation (in parentheses) Reading Times on the Target Word in Each Condition

Group	Condition	First fixation duration (ms)	Single fixation duration (ms)	Gaze duration (ms)	Selective regression path duration (ms)	Total reading time (ms)
Adults	Identity	211 (71)	212 (71)	239 (101)	261 (128)	356 (261)
	P+ O-	228 (80)	237 (81)	264 (100)	284 (117)	387 (246)
	P- O-	233 (76)	241 (78)	275 (116)	301 (130)	387 (229)
Children	Identity	273 (120)	283 (116)	410 (298)	498 (353)	682 (597)
	P+ O-	292 (140)	307 (132)	434 (303)	512 (359)	665 (504)
	P- O-	279 (134)	300 (135)	447 (324)	539 (352)	727 (541)

Two lme models were run for each dependent measure. Model 1 compared the letter substitution previews (P+ O-, P- O-) to the identity condition, with participant group included as an interaction term. This allowed us to examine the potential costs associated with a nonword preview, examining whether the participants displayed preview benefit, with the adults acting as the baseline. Then contrasts (second lme model) were run to directly compare the letter substitution preview conditions, in order to determine whether phonology might play a role in the first-letter bias. Effects were considered significant when $|t| > 1.96$.

As word length varied (stimuli word length ranged between 5-7 letters), lme models were also run with length as a factor. For all of the dependent measures, word length had no significant effect (intercepts only models; *t*s < +/-1.54). Formal model comparisons were also conducted to examine word length's role within our data. The comparisons showed, again within all dependent measures, that including word length did not improve the fit of our models and contrasts, *p*s > .242, thus, we report the results from the models that do not include word length for the sake of brevity and simplicity. In addition, given that each participant was presented with three different previews of each target word (six files accounted for every combination possible), formal model comparisons were conducted to determine whether preview presentation order might have had an effect on participants' processing (reading times would be expected to decrease on any given target word over the second and third presentations of that target, akin to a practice effect). The comparisons showed that for first fixation duration, single fixation duration, and gaze duration, the inclusion of presentation did not improve the fit of our models and contrasts, *p*s > .104. Within selective regression path duration and total reading time, however, presentation did improve model fit when included as a main effect (additive) term, *p*s < .001.[29] The effect of presentation is considered and summarised here, as the findings are not pertinent to the interpretation of our experimental manipulations: reading times were significantly faster after the second and third time a target word was presented to participants, relative to the first time (as shown in the model and contrast results reported below- see also Appendix B; B.3; Figure B.3.1), but reading times were not significantly different between the second and third times that participants saw each target word (see Appendix B; B.3; Table B.3.3 and Figure B.3.1).

To reiterate, Model 1 used the identity condition as a baseline, with each of the substituted letter preview conditions (P+ O-, P- O-) compared to it, and with the children's

---

[29] Given the inclusion of presentation as an additive term (vs. no term) significantly improved model fit within these measures, we also ran model comparisons comparing its inclusion as an additive term against its inclusion as an interactive term. The models with presentation included as an additive term were a better fit than the models with presentation included as an interactive term, *p*s > .278.

data compared to the adult data. The results of this model, for each dependent measure, are shown in Tables 5.5 (first fixation duration, single fixation duration, and gaze duration) and 5.6 (selective regression path duration and total reading time; note that the models for these measures also includes presentation order). The contrasts directly compared the P+ O- previews to the P- O- previews, in order to examine the effect the first letter's phonological code being maintained in preview had on both adults' and children's parafoveal pre-processing (i.e., it could be determined whether the first letter's phonology being preserved in preview facilitated lexical identification compared to when it was disrupted in preview).

Table 5.5

Output from Model 1, and the Contrasts, for First Fixation Duration, Single Fixation Duration, and Gaze Duration

	First fixation duration				Single fixation duration				Gaze duration			
	b	SE	*t*	*p*	b	SE	*t*	*p*	b	SE	*t*	*p*
Adults, Identity (Int)	5.30	.02	217.87	< .001	5.31	.03	189.11	< .001	5.40	.04	133.90	< .001
Adults, Children	**.23**	**.03**	**7.16**	**< .001**	**.28**	**.04**	**7.25**	**< .001**	**.45**	**.05**	**8.43**	**< .001**
Adults, P+ O-	**.08**	**.02**	**3.09**	**.003**	**.11**	**.02**	**5.11**	**< .001**	**.11**	**.02**	**4.57**	**< .001**
Adults, P- O-	**.10**	**.02**	**4.70**	**< .001**	**.13**	**.02**	**6.01**	**< .001**	**.14**	**.02**	**6.02**	**< .001**
Children × P+ O-	-.03	.03	-.72	.471	-.04	.03	-1.28	.199	-.04	.03	-1.25	.210
Children × P- O-	**-.10**	**.03**	**-3.29**	**.002**	**-.09**	**.03**	**-2.86**	**.004**	**-.07**	**.03**	**-2.14**	**.032**
Contrasts												
Intercept	5.45	.02	271.16	< .001	5.50	.02	236.10	< .001	5.69	.04	157.41	< .001
Adults, P+ O- vs. P- O-	-.02	.02	-1.06	.288	-.02	.02	-.89	.372	-.03	.02	-1.46	.146
Children, P+ O- vs. P- O-	**.05**	**.02**	**2.47**	**.014**	.03	.03	1.32	.188	-.01	.02	-.28	.778

Note. The reading time data were log transformed prior to analysis, so the model estimates cannot be directly interpreted. Significant effects are marked in bold. Following trimming, the syntax for first fixation duration was: *depvar ~ Group * condition + (1 + condition|Participant) + (1|targetno)*, and for single fixation duration and gaze duration the syntax, as intercepts only models, was: *depvar ~ Group * condition + (1|Participant) + (1|targetno)*. Within the contrasts, after trimming, the syntax for all measures, as intercepts only models, was: *depvar ~ GroupByCond + (1|Participant) + (1|targetno)*.

Table 5.6

Output from Model 1, and the Contrasts, for Selective Regression Path Duration and Total Reading Time

	Selective regression path duration				Total reading time			
	b	SE	*t*	*p*	b	SE	*t*	*p*
Adults, Identity, Presentation 1 (Int)	5.54	.05	111.63	< .001	5.80	.06	101.84	< .001
Group (Adults vs. Children)	**.58**	**.06**	**9.03**	**< .001**	**.60**	**.07**	**8.21**	**< .001**
Adults, P+ O-	**.10**	**.03**	**3.88**	**< .001**	**.11**	**.03**	**4.16**	**< .001**
Adults, P- O-	**.16**	**.02**	**6.54**	**< .001**	**.13**	**.03**	**4.82**	**< .001**
Adults, Presentation 2	**-.11**	**.02**	**-6.35**	**< .001**	**-.16**	**.02**	**-8.21**	**< .001**
Adults, Presentation 3	**-.11**	**.02**	**-6.03**	**< .001**	**-.16**	**.02**	**-8.15**	**< .001**
Children × P+ O-	-.07	.04	-1.78	.079	**-.11**	**.04**	**-2.75**	**.006**
Children × P- O-	**-.07**	**.03**	**-2.12**	**.037**	-.05	.04	-1.20	.230
Contrasts								
Intercept, Presentation 1	5.90	.04	131.52	< .001	6.15	.05	119.46	< .001
Adults, P+ O- vs. P- O-	**-.06**	**.02**	**-2.58**	**.010**	-.02	.03	-.66	.508
Children, P+ O- vs. P- O-	**-.05**	**.02**	**-2.33**	**.020**	**-.08**	**.03**	**-2.94**	**.003**
Presentation 2	**-.12**	**.02**	**-7.51**	**< .001**	**-.16**	**.02**	**-8.17**	**< .001**
Presentation 3	**-.11**	**.02**	**-6.88**	**< .001**	**-.16**	**.02**	**-8.15**	**< .001**

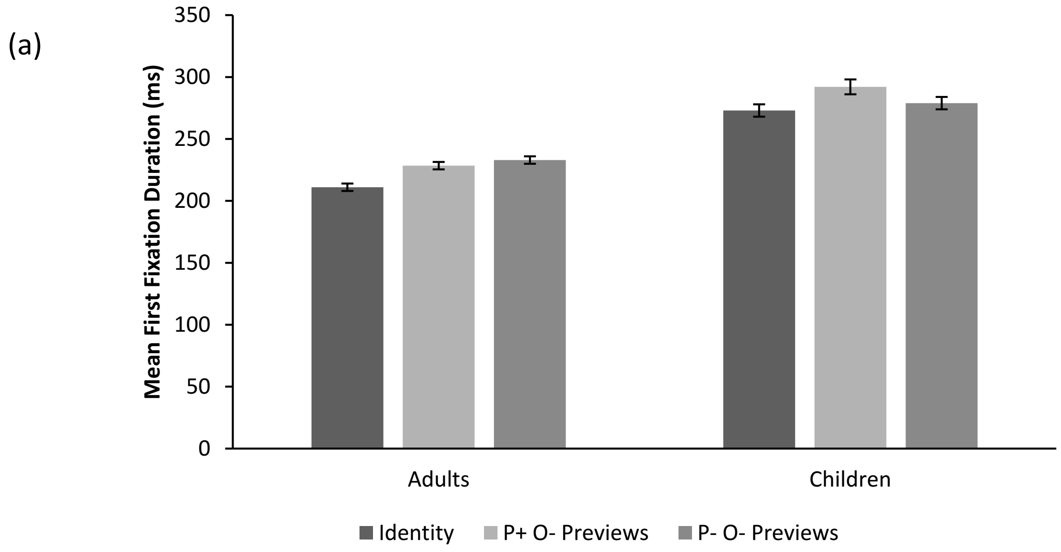
Note. The reading time data were log transformed prior to analysis, so the model estimates cannot be directly interpreted. Significant effects are

marked in bold. Following trimming, the syntax for selective regression path duration was: *depvar ~ Group * condition + presentation + (1 + condition|Participant) + (1|targetno)*, and the syntax for total reading time, as an intercepts only model, was: *depvar ~ Group * condition + presentation + (1|Participant) + (1|targetno)*. Within the contrasts, after trimming, the syntax for both measures, as intercepts only models, was: *depvar ~ GroupByCond + presentation + (1|Participant) + (1|targetno)*.

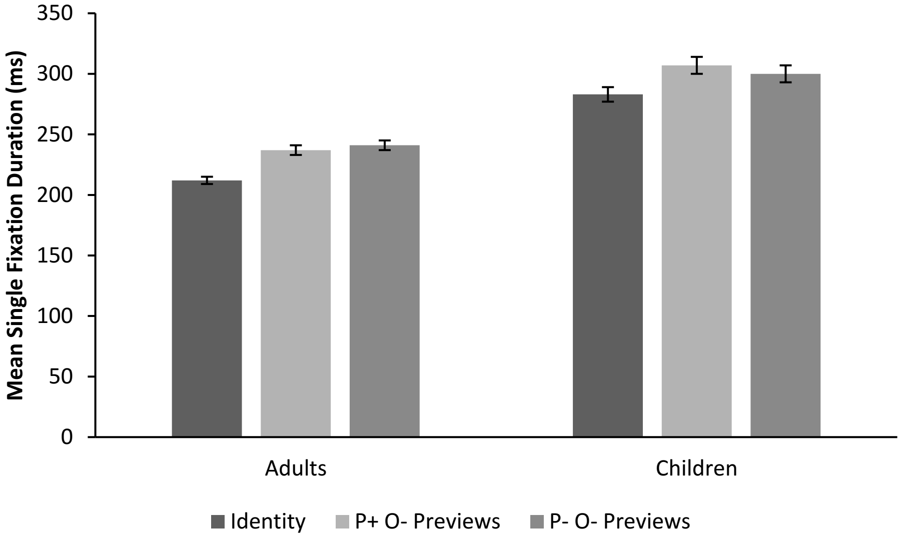
Firstly, within all measures, there were significant group differences: the children displayed significantly longer reading times than the adults (see Tables 5.4, 5.5, and 5.6, and Figure 5.2). Second, across all measures, the adults displayed clear preview benefit, such that both substituted letter previews resulted in longer reading times than the identity preview.

Figure 5.2

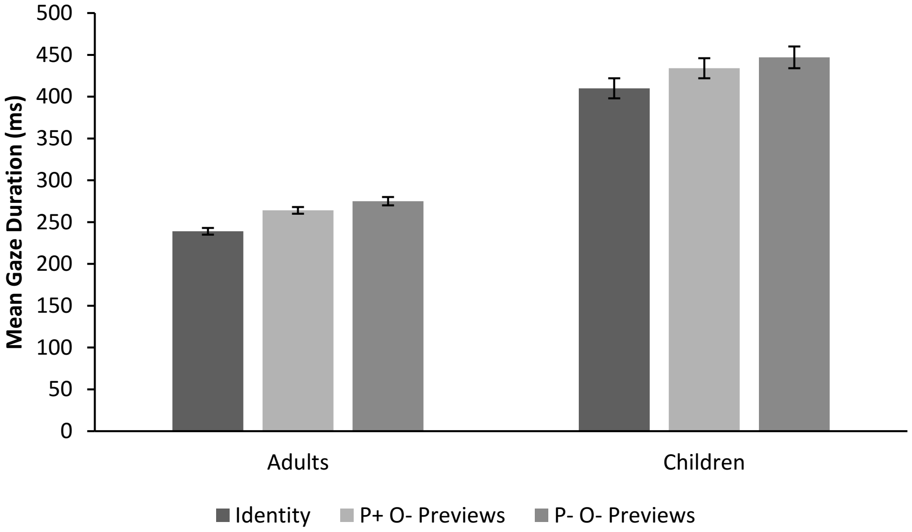
Mean First Fixation Durations (a), Single Fixation Durations (b), Gaze Durations (c), Selective Regression Path Durations (d), and Total Reading Times (e) on Identity, P+ O- Previews, and P- O- Previews for Both Adults and Children.



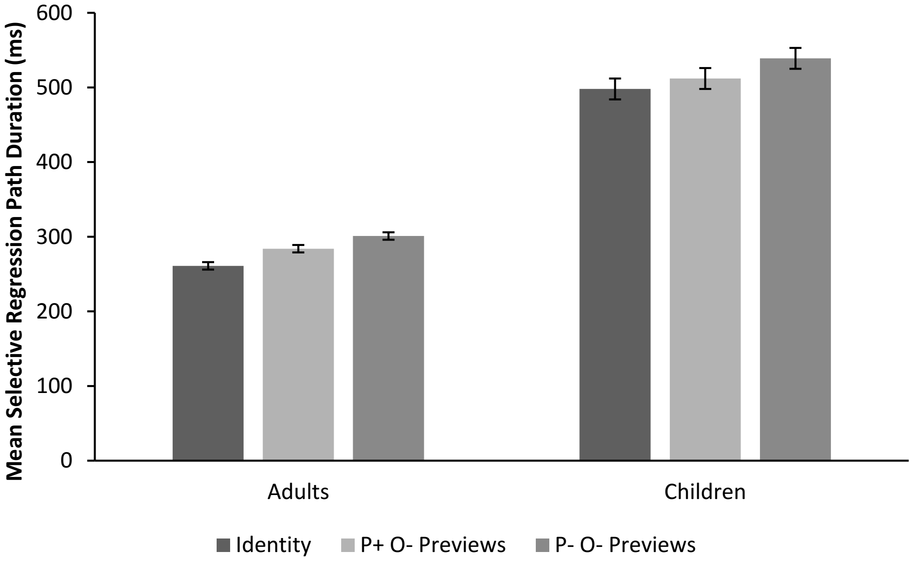
(b)

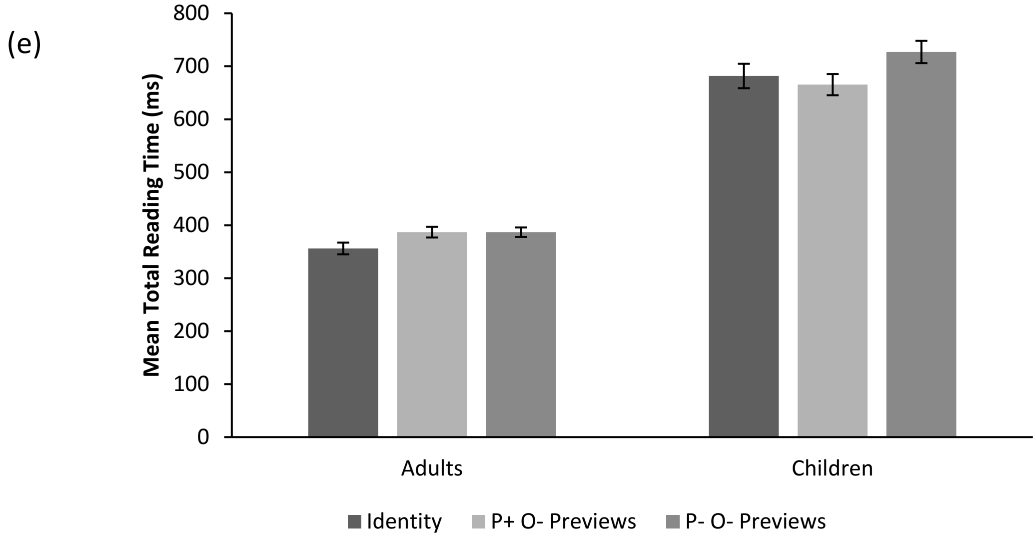

Mean Single Fixation Duration (ms)
350
300
250
200
150
100
50
0
Adults
Children
Identity
P+ O- Previews
P- O- Previews

(c)


Mean Gaze Duration (ms)
500
450
400
350
300
250
200
150
100
50
0
Adults
Children
Identity
P+ O- Previews
P- O- Previews

(d)


Mean Selective Regression Path Duration (ms)
600
500
400
300
200
100
0
Adults
Children
Identity
P+ O- Previews
P- O- Previews



The children's data were largely consistent with the adult data; reading times were disrupted after P+ O- previews relative to the identity condition (with the exception of total reading time). With respect to the second substituted letter condition, there were some differences (see Tables 5.5 and 5.6). Whilst numerical differences show that both groups displayed longer reading times following P- O- previews than identity previews (see Table 5.4), the magnitude of the effect was smaller in the children's data for early measures of processing (reflected in the significant interaction terms for "Children × P- O-").

Given the multiple comparisons undertaken within Model 1, we also ran Model 1 (without the intercept) for each dependent measure using the *glht* function (package multcomp) to adjust our *p* values for multiple comparisons (Hothorn et al., 2008). Two effects went from being significant to non-significant after using this correction technique: within gaze duration and selective regression path duration the interactions "Children × P- O-" became non-significant (*p* = .125 and *p* = .184, respectively). These analyses, therefore, also show that both adults and children displayed longer reading times after both letter substitution previews compared to an identity preview.

Critically, the contrasts show that adults did not benefit from the first letter's phonology being preserved in preview within early processing (they displayed similar reading times between the P+ O- previews and the P- O- previews; only in the later measure of selective regression path duration was there any difference between these two substitution conditions). Similarly, the children did not show any advantage from the preservation of the first letter's phonology in early measures of processing (see Tables 5.4 and 5.5, and Figure 5.2). In the later measures of selective regression path duration and total reading time, though, the children did benefit from the first letter's phonology being maintained in preview: they displayed faster reading times after P+ O- previews compared to P- O- previews (see Tables 5.4 and 5.6, and Figure 5.2).

Again, given the multiple comparisons undertaken within the key contrasts, we also ran the contrasts (without the intercept) for each dependent measure where there was a significant result present using the *glht* function (package multcomp) to adjust our *p* values (Hothorn et al., 2008). In this analysis, the difference between the two letter substitution conditions in children's selective regression path duration approached significance (*p* = .075). Critically, the benefit from phonology in preview was absent within early measures of processing; the benefit was consistently observed in total reading time.

5.4.2.1 Bayesian Analyses

Of critical interest within our results was the null effect of the first letter's phonology being preserved in preview (the comparison of the two letter substitution preview conditions). Consequently, Bayesian analyses were conducted to assess the strength of the evidence for the null and alternative hypotheses, wherever null effects were present within the contrasts. The analyses were conducted using the BayesFactor package (Morey & Rouder, 2013), for the Cauchy priors on effect size the default scale value (.5) was used, and 100,000 Monte Carlo iterations were specified. A low Bayes factor (< 1) indicates evidence for the null hypothesis and a high Bayes factor (> 1) provides evidence for the alternative hypothesis. For all models/contrasts items and subjects were specified as random factors.

Within the contrasts we examined any null effects that were present within each measure by comparing a specified model, which coded the two experimental preview conditions (P+ O- and P- O-) separately for the adults and the children (*PhonAdults*/*PhonChildren*), against the default intercept only model. The Bayes factors from the analyses were .16 for the adults in first fixation duration, .14 for both the adults and the children in single fixation duration, .27 for the adults and .07 for the children in gaze duration, and .07 for the adults in total reading time. Using the commonly cited evidence categories for Bayes factors, where a Bayes factor < .33 provides substantial evidence for a null effect, and a Bayes factor < .10 provides strong evidence, these Bayesian analyses indicate substantial evidence (and in the case of children's gaze durations and adults' total reading times, strong evidence) for the null hypothesis (i.e., the adults and the children were not gaining a significant benefit from the first letter's phonology being maintained in preview).

We also conducted Bayesian analyses on the null interactions within Model 1, in order to determine whether the children were indeed patterning like the adults. These null interactions were examined by comparing a model that specified the fixed factors of group and condition (e.g., *Group* + *condition*) with a model that additionally contained an interaction term (e.g., *Group* + *condition* + *Group:condition*). The Bayesian analyses indicated substantial or strong evidence for the null hypothesis regarding the interactive term between group and P+ O- previews (.09 for first fixation duration, .15 for single fixation duration, .12 for gaze duration, and .49 for selective regression path duration), and substantial evidence for the null hypothesis regarding the interactive term between group and P- O- previews in total reading time (Bayes factor = .14). This suggests that, overall, the children's parafoveal pre-processing of these preview conditions was consistent with the adults' pre-processing within these measures (i.e., like the adults, the children were displaying a cost from substituted letter previews compared to an identity preview).

5.5 Discussion

We investigated the first-letter bias within parafoveal pre-processing, examining what drives this effect during silent sentence reading: orthographic or phonological encoding. We compared the effects of our manipulations in skilled adult and beginner child readers. Firstly, as predicted, we found significant group differences: the children displayed significantly longer reading times than the adults, consistent with past research (e.g., Blythe & Joseph, 2011). The children's rate of lexical processing during reading was slower, and less efficient, than that of the adults, consistent with simulations of adults' and children's eye movement behaviour during reading within the E-Z Reader model (Mancheva et al., 2015; Reichle et al., 2013).

Nevertheless, as predicted, both the adults and the children displayed a first-letter bias: when the first letter was substituted in preview, compared to the identity condition, this disrupted their ability to lexically identify word *n*+1 (consistent with past research; e.g., Milledge et al., 2020). Moreover, as predicted, comparison of the two experimental preview conditions elucidated the cause of the first-letter bias. Within both adults' and children's first-pass reading, there was no evidence of the first-letter bias being phonologically driven; rather, the data are indicative of, primarily, orthography driving the importance of the first letter in preview. This may seem, at first glance, to contradict past research that has shown that skilled adult and beginner child readers process phonological information from word *n*+1 (e.g., Milledge et al., 2021a; Pollatsek et al., 1992); however, upon closer inspection of the experimental manipulations and patterns of effects, it seems that any benefit from phonology in preview is dependent upon access to the correct orthographic code of the first letter.

Research findings are consistent with this idea of the first letter playing a vital role in facilitating readers' ability to benefit from phonology in preview. Pollatsek et al. (1992) found that adult readers, on average, did not display as much benefit from homophone previews over spelling control previews when the first letter of a target word was substituted in preview (e.g., *c* substituted with *s* in preview; *shoot* vs. *shout* as previews for *chute*), in comparison to when the first letter was maintained in preview (e.g., *beech* vs. *bench* as previews for *beach*)- 20 ms benefit vs. 37 ms benefit in first fixation duration, respectively (Experiment 2). Similar effects have also been found within children. Milledge et al. (2021a) found that, especially within the early processing of their orthographically dissimilar previews (half of these previews involved the substitution of at least the first letter), the children did not gain a benefit from intact phonology in preview. The children displayed longer reading times on pseudohomophone previews than spelling control previews (e.g., *kley* vs. *bloy* as previews for *clay*). It would seem, therefore, that preserving the orthographic code of the first letter of word *n*+1 in preview facilitates the efficient extraction of phonological information from that word for both adults and children.

Indeed, within the present research, we found no differences between reading times on target words after P+ O- previews compared to P- O- previews within the adult readers' early processing; indicating that the first-letter bias was, primarily, orthographically, not phonologically, driven. Regarding the children, they even displayed longer reading times in early processing (first fixation duration) on P+ O- previews than P- O- previews. Strikingly, when incorrect orthographic information was present for the first letter in preview, the children were unable to benefit from the first letter's phonology being present in preview; in fact, they suffered a cost. Both the adults and the children were unable to efficiently make use of the correct phonological information of the first letter being present in preview due to the orthographic manipulation of that letter, with this effect especially evident within the children. As such, the preservation of the orthographic code of the first letter would appear to be critical to both adult and, potentially especially, child readers' parafoveal pre-processing and early lexical identification processes within English, broadly consistent with past research (Milledge et al., 2021a; Pollatsek et al., 1992).

Within both adult and child readers, the first letter's orthographic code would appear to be activated first, followed by its phonological code (e.g., Grainger, Dufau, et al., 2016). Given the notion that, within the lexicon, orthographic lexical representations activate phonological lexical representations (e.g., Perry et al., 2007), when the adult and child readers had, for example, *sircle* as a preview for *circle*, within the lexicon orthographic lexical representations for word *n*+1 would have been, incorrectly, activated for words beginning with *s*. Consequently, despite intact phonological information being present, the presence of incorrect orthographic information (having the first letter *s* in preview rather than *c*) caused an immediate cost to both the adults' and the children's ability to lexically identify word *n*+1. This could have been further compounded by the nature of the English language and its inconsistent GPCs: for example, *c* can have a /s/ sound or a /k/ sound and *g* can have a /j/ sound or a /g/ sound. When readers came to directly fixate the correct target word, in addition to the subsequent need to activate correct orthographic lexical representations (e.g., representations for words beginning with *c* rather than *s*), this would have resulted in the activation of multiple phonological lexical representations, given the first letter substitutions made within the P+ O- previews had more than one sound associated with them. Thus, readers would have been faced with an increasing number of competing lexical representations. Essentially, the unpredictable and complex nature of English (e.g., Schmalz et al., 2015) could have caused extra processing costs for the readers, with the first letter driving this cost.

Overall, the present findings regarding a first-letter bias being present within both skilled adult and beginner child readers, and the primarily orthographically driven nature of this bias, are consistent with models of orthographic encoding (e.g., Grainger & Ziegler, 2011; Spatial Coding model, Davis, 2010b; SERIOL, Whitney, 2001). We note again, though, that these models relate to isolated word identification under direct fixation and can, therefore, only make inferences about how this might extend to processing of word *n*+1 within natural sentence reading. Nonetheless, the early orthographic nature of the first-letter bias found is consistent with the SERIOL (Whitney, 2001) and Spatial Coding (Davis, 2010b) models of visual word recognition. Both models posit that the first letter plays an important role in word identification processes, given sequential processing of letters within a word (Whitney, 2001) and dynamic end-letter marking (Davis, 2010b); within both models the first letter receives increased activation/weight.

Regarding how these results relate to Grainger and Ziegler's (2011) model, the adults and the children appeared to display similar processing within their fine-grained routes (as previously found by Milledge et al., 2021a). Both the adults and the children displayed an immediate cost when phonology was maintained in preview, requiring some form of sublexical conversions of print-to-sound to be undertaken for word *n*+1, but orthography was manipulated (P+ O- previews). This is as would be expected given the fine-grained route's limited flexibility with regard to orthographic encoding. Within both the skilled adult and beginner child readers' early processing within the fine-grained route, the presence of the first letter's correct orthographic code would appear to be key to the orthographic encoding that takes place in order to achieve lexical identification, with the correct orthographic code enabling effective, and efficient, processing of phonology. The late occurrence of the benefit from the first letter's phonology in preview highlights the inefficiency with which phonological information could be extracted from word *n*+1 by readers of English when incorrect orthographic information was present in preview.

We also found differences, as predicted, in the time course of parafoveal pre-processing between adult and child readers. The adults and the children differed in the time course of their processing of P- O- previews (where both phonology and orthography were manipulated); although the children showed numerical costs, they were less affected than the adults, within early processing, by these previews compared to the identity previews. Within Grainger and Ziegler's (2011) model, the coarse-grained route would have been used for these previews, which allows more flexibility with regard to orthographic encoding. This flexibility could be increased by children's orthographic representations being encoded with less precision compared to those of adults (e.g., Perfetti, 2007). The adults, with their more precisely encoded orthographic representations, would be more reliant on whole-word orthography in preview (as provided by the identity previews); whilst for children, if orthographic forms are less precisely encoded, less of an immediate cost would be expected when orthography is manipulated in preview, within the coarse-grained route. Broadly consistent with the findings of Milledge et al. (2020), this is suggestive of developmental change within the tuning of orthographic processing (e.g., Castles et al., 2007). Moreover, this suggests that 8- to 9-year-old child readers of English might still be developing their coarse-grained routes of processing (Grainger & Ziegler, 2011), with, presumably, this development continuing over time, as beginner readers progress to be skilled readers and develop higher quality lexical representations (e.g., Perfetti, 2007).

In conclusion, the present experiment provides novel evidence of the first-letter bias in parafoveal pre-processing being orthographically driven for both adults and children. Moreover, this experiment also provides novel insight into the time course of both adults' and children's ability to extract phonological information from the first letter of a word in preview when its orthography is manipulated. Of note, overall, is the critical role the first letter's orthography plays in preview, facilitating both adults' and children's efficient- early- processing of phonology in English.

Chapter 6 General Discussion

6.1 Summary of Experimental Research and Findings

Previous research on parafoveal pre-processing has predominantly focused on skilled adult readers. I have presented three empirical papers within this thesis that have furthered our understanding of this skill, so crucial to skilled adult reading (e.g., Rayner, Liversedge, et al., 2006), within beginner child readers of English, using the boundary paradigm (Rayner, 1975).

As delineated in Chapter 1, there were clear gaps in knowledge regarding parafoveal pre-processing within child readers of English; namely, whether there is developmental change in the time course of this processing, and what characteristics children extract from word *n*+1. As such, firstly, by comparing children's parafoveal pre-processing to that of adults, I was able to examine whether processing difficulty leads to differences in the time course of skilled adult and beginner child readers' parafoveal pre-processing. As outlined in Chapter 1, the time course of lexical processing differs between adult and child readers within foveal processing- lexical processing is slower in children (e.g., Blythe et al., 2006, 2009, 2011, 2015; Joseph et al., 2009; Pagán et al., 2021; Reichle et al., 2013; Tiffin-Richards & Schroeder, 2015b). It was thought that this reduced rate of lexical processing (i.e., the increased processing difficulty) within children, compared to adults, could lead to differences in the time course of their parafoveal pre-processing. Secondly, I was able to examine what characteristics of word *n*+1 skilled adult and beginner child readers extract, in order to facilitate their lexical identification. Specifically, within this thesis, the roles orthographic and phonological information play in preview was examined.

Orthography and phonology are intrinsically linked within English: the printed form of a word conveys, and allows access to, phonological information. In Chapter 2

(Milledge & Blythe, 2019), I outlined how current theories of reading development do not account for the important role phonological processing plays in lexical identification throughout reading development; given they propose an increasing reliance on orthographic processing, whilst phonological processing becomes less important (e.g., Ehri, 1995, 1998, 1999, 2005, 2007; Frith, 1985). Indeed, counter to this supposition, children have been found to continue to use phonology after they have stopped using overt decoding strategies (e.g., Blythe et al., 2015). Overall, the literature clearly supports the notion of developmental change in the nature of phonological processing. The ability to extract phonological information from preview is a core aspect of this developmental change in phonological processing.

Skilled adult readers of English have been shown to extract phonological information from word *n*+1, and this processing is contingent on their extraction of orthographic information from word *n*+1 (Pollatsek et al., 1992). Consequently, prior to determining whether child readers of English also display this developmental change in their phonological processing (i.e., processing of phonology from preview), it was necessary to determine whether children are sensitive to information across the whole-word form of word *n*+1. Then, the role phonology plays in preview could be examined. After addressing these gaps in knowledge, the interplay between orthographic and phonological information with regard to children’s processing of word *n*+1 could start to be examined.

In Chapter 3 (Milledge et al., 2020), therefore, I examined outstanding questions regarding children’s parafoveal pre-processing of orthography: the extent to which child readers are able to extract information from whole-word forms (up to six letters) in preview, and whether external letters in preview are more facilitative to lexical identification than internal letters in preview. Child readers were able to extract orthographic information from an entire word in preview; their processing was disrupted (shown by longer reading times) when the end letter of a target word was substituted in preview. For both the adult and child readers, the external letters being substituted in preview was more harmful to their ability to lexically identify word *n*+1 than the internal letters being substituted in preview. The external letters of a word are, therefore, crucial for both adults' and children's lexical identification processes. Moreover, within the external letters, it was found that the first letter played an important role within both adults' and children's processing: when the first letter was substituted in preview this came at an immediate, and similar, cost to both groups of readers. Interestingly, apart from when the first letter was substituted in preview, the children's parafoveal pre-processing of orthography was slower compared to that of the adults.

In Chapter 4 (Milledge et al., 2021a), I subsequently examined whether child readers can extract phonological information from word *n*+1, and the extent to which this is constrained by orthography. Both adult and child readers were found to only display an advantage (i.e., faster reading times) from phonology being preserved in preview, relative to a spelling control, when the previews were orthographically similar. When the previews were orthographically dissimilar, however, no advantage from phonology, over a spelling control, was found. Of note is that within the orthographically dissimilar stimuli, in contrast to the orthographically similar stimuli, for half of these stimuli, at least one of the letters that was substituted in preview was the first letter. This substitution of the first letter in preview harmed the readers' ability to extract, and benefit from, the intact phonological information being present for word *n*+1. Overall, this research indicates, similar to Milledge et al. (2020), that 8- to 9-year-old readers extract similar characteristics from word *n*+1 to skilled adult readers.

Following on from the results of the experiments reported in the previous chapters, in Chapter 5 (Milledge et al., 2021b) I examined whether the importance of the first letter (first-letter bias) in preview is phonologically or orthographically driven. Both the adults' and the children's ability to lexically identify word *n*+1 was disrupted when the first letter was manipulated in preview; they displayed a first-letter bias. Moreover, there was no benefit seen in adults' and children's first-pass reading times from the first letter's phonology being maintained in preview. This suggests that the first-letter bias was primarily orthographically driven. Thus, preserving the orthographic code of the first letter in preview is key to both adults' and children's ability to efficiently process phonology. In addition, it was found that the children's parafoveal pre-processing of orthography was slower compared to that of the adults. The results also provide further evidence of child readers extracting the same information from word *n*+1 as adult readers.

Summarising what was found with regard to children's parafoveal pre-processing in English, the key results of the experiments are as follows: 1) children are able to extract orthographic information from the whole of word *n*+1 (up to six letters); 2) children extract phonological information from preview, dependent on orthography; 3) children show a first-letter bias; 4) this first-letter bias is an orthographic effect; 5) their parafoveal pre-processing is slower than that of adults, typically in the order of one fixation.

Before further in-depth discussion of these results, I first address how these findings might be considered in respect of the type of parafoveal masking used. That is, the use of parafoveal letter mask previews instead of degraded previews.

6.2 Parafoveal Mask Considerations

The boundary paradigm (Rayner, 1975) was used in all of the experiments reported within this thesis. As highlighted in Chapter 1 (pp. 27-29), consideration is needed over the interpretation of effects in relation to the type of parafoveal mask that is used, due to the argument that parafoveal letter masks induce extra processing costs within child readers, leading to preview benefit being overestimated (e.g., Marx et al., 2015).

Of note, even if the use of parafoveal letter masks (i.e., letter substitutions in preview) does lead to overestimation of preview benefit within the silent sentence reading of children, this does not fundamentally change the results of the experiments reported within this thesis. For example, a key finding was that there were differences in the time course of adults' and children's parafoveal pre-processing, such that the extraction of orthographic information from word *n*+1 was slower within children compared to that of the adults (i.e., the preview benefit effect was slower to appear). The concern over the overestimation of preview benefit through the use of letter substitution masks refers to the magnitude of the effect, not the time course. Ultimately, then, this would not change the important finding that parafoveal pre-processing of orthographic information (letter substitutions) was slower in children compared to adults.

In addition, Marx et al.'s (2015) data suggests that parafoveal 'x' masks caused more interference to the children's processing than the parafoveal letter mask previews. Within all of the experiments reported within this thesis, parafoveal letter mask previews were used. Interestingly, Vasilev and Angele (2017) found that the less 'word-like' a preview was, the larger the effect of preview benefit. The vast majority of the previews I used were orthographically legal and pronounceable, more 'word-like', in comparison to the parafoveal letter mask previews used by Marx et al. (2015); e.g., *Zwmn* as a preview for *Haus* vs. *romlun* as a preview for *sister* (Milledge et al., 2020). This suggests that any overestimation of preview benefit would have been limited, to an extent, within the present research.

Moreover, even if the absolute cost of letter substitutions led to preview benefit being overestimated, critically, this does not detract from the findings regarding the relative costs/benefits of letter substitutions in preview. That is, the costs/benefits found when comparing the experimental previews (e.g., Chapter 3- external letter vs. internal letter substitutions in preview; Chapter 4- pseudohomophones previews vs. spelling control previews within the orthographically similar stimuli, etc.). Consequently, the use of parafoveal letter mask previews (letter substitutions) within the experiments reported in this thesis should not detract from the results found.

6.3 Developmental Change in Parafoveal Pre-processing

The results of the experiments reported in this thesis provide evidence of developmental change in parafoveal pre-processing, especially with respect to the extraction of orthographic information from word *n*+1. Recall that children's lexical processing is thought to be slower than that of adults, as indexed by overall group differences in reading times (e.g., Blythe et al., 2009, 2011; Joseph et al., 2009; Pagán et al., 2021), with further support for this argument coming from simulations within a computational model of eye movement control during reading (Mancheva et al., 2015; Reichle et al., 2013). Given children's slower lexical processing, then, it is unsurprising that delays were found in the time course of children's parafoveal pre-processing relative to that of the adults in the present set of experiments. If children are slower to lexically process word *n*, it stands to reason that they will also be slower to extract information from word *n*+1. As a result, this not only inflates reading times across entire sentences (i.e., total sentence reading times), but means that each word in a sentence is processed at a slower rate when it is fixated by a child compared to an adult (see Häikiö et al., 2010; Pagán et al., 2016; Tiffin-Richards & Schroeder, 2015a). This is non-trivial, given that research has shown that children require a similar amount of time to adults to extract visual information during direct fixation (Blythe et al., 2009). Longer reading times in children must, therefore, be indicative of slower lexical processing in children compared to adults.

The results regarding the role lexical processing plays in the developmental change in eye movement behaviour, rather than differences in visual encoding (Blythe et al., 2009) and oculomotor control (e.g., Reichle et al., 2013), not only apply to foveal processing but, as the present research shows, also extends to parafoveal pre-processing. For example, within Chapter 3 it was found that, in addition to the children's overall reading times on target words being longer than those of the adults (indicating slower lexical processing), children, like adults, were sensitive to end letter substitutions in preview. Consequently, visual factors, at least concerning words of six letters long, were not constraining their parafoveal pre-processing of word *n*+1. As such, improvements in reading, that typically occur with development, reflect increased speed in lexical processing. This is also demonstrated through how the perceptual/letter identity span within both adult and child readers is constrained by processing demands. The more efficient foveal lexical processing becomes, with age and reading skill (e.g., Häikiö et al., 2009; Rayner, 1986; Sperlich et al., 2015; Veldre & Andrews, 2014), the more resources can be allocated to upcoming information in preview. In essence, it would indeed appear that children's more resource-demanding, and slower, lexical processing of word *n* (e.g., Blythe et al., 2009, 2011, 2015; Joseph et al., 2009, 2013; Mancheva et al., 2015; Reichle et al., 2013; Tiffin-Richards & Schroeder, 2015b), leads to fewer resources being available to start processing of word *n*+1, resulting in slower parafoveal pre-processing compared to that of adults. Specifically, within the present research, the children displayed a slower rate of orthographic pre-processing, generally in the order of one fixation.

Insight into the processing time course differences found within orthographic encoding between the adults and the children can be gained through simulations that were run using the E-Z Reader model. Mancheva et al. (2015) found that orthographic knowledge, rather than phonological processing skill, accounted for variance in children's eye movement measures and variability in reading skill (Simulation 2). Consequently, overall, not only do differences in lexical processing seem to drive the developmental differences between adults' and children's eye movements during reading, but orthographic processing has a marked contribution to the development of skilled lexical identification processes.

6.3.1 Lexical Quality Hypothesis and Reading Skill

In this way, differences in the quality of lexical representations (e.g., Perfetti, 2007) could play a role in the developmental differences found between skilled adult and beginner child readers' eye movements during reading. Lexical representations could be less precisely encoded within children, relative to those of adults, resulting in slower lexical identification. This difference in the speed of lexical processing is demonstrated by consistently longer reading times being found in child readers, compared to adults, within both the present research and past research (e.g., Blythe et al., 2006, 2009, 2011). Skilled readers, in contrast to less skilled/beginner readers, have a greater number of high quality lexical representations.

Consistent with this idea, compared to skilled adult readers, the children displayed slower parafoveal pre-processing; specifically, slower extraction of orthographic information from word *n*+1. When orthographic information was manipulated in preview, skilled adult readers experienced more immediate costs to their processing. The adult readers were less able to access lexical representation candidates based on partial orthographic information from preview than beginner child readers during first-pass reading. The higher quality of the adults' lexical representations meant that they had a greater reliance on correct orthographic information being present.

Interestingly, the quality of orthographic lexical representations depends on reading ability. Pagán et al. (2021) found, within a reading-like task, that reading ability modulated orthographic processing within child readers. The eye movements of children with higher reading ability behaved in a manner more consistent with the skilled adult readers (i.e., their processing of orthography was more similar to that of the adults). This suggests that their lexical processing was faster (and more adult-like) than the children with lower reading ability, driven by higher quality orthographic lexical representations. As beginner readers develop higher quality lexical representations, in conjunction with development in reading skill, their eye movement behaviour becomes more adult-like, due to increases in the speed of their lexical processing (Luke et al., 2015; Mancheva et al., 2015; Pagán et al., 2021; Reichle et al., 2013).

Overall, again, both theory and research highlight the importance of orthography in facilitating lexical identification. Past research has shown, with regard to parafoveal pre-processing, that the letter identity span is modulated by reading skill: more skilled child readers have larger letter identity spans than less skilled child readers (Häikiö et al., 2009). As such, within children, differences in reading skill affect the spatial extent of the orthographic encoding they are able to undertake with regard to word *n*+1. The research reported within this thesis found that child readers were undertaking similar orthographic encoding within preview to that of adult readers, but it is of note that all of the child participants were, at least, "average" readers for their age (Wechsler, 2005). Potentially, less skilled child readers have a reduced ability to extract orthographic information from word *n*+1, in comparison to more skilled child readers, given the lower quality of their orthographic lexical representations (e.g., Mancheva et al., 2015; Pagán et al., 2021), and constraints as to the spatial extent of the area from which they can parafoveally pre-process information (Häikiö et al., 2009).

To reiterate, speed of lexical processing, indexing the ease with which lexical identification occurs, is slower in child readers compared to adult readers, within both foveal (e.g., Blythe et al., 2009, 2011, 2015; Joseph et al., 2009; Tiffin-Richards & Schroeder, 2015b) and parafoveal (Häikiö et al., 2010; Milledge et al., 2020, 2021a,b; Pagán et al., 2016) processing. This reduced rate of lexical processing is indicative of differences in reading skill. Indeed, within the research reported within this thesis (Milledge et al., 2020, 2021a,b), the children were found for their age to be less skilled readers than the adults (as measured by their composite standardised scores on the WIAT-II UK; Wechsler, 2005), although they were at least "average" readers for their age. This is unsurprising, given that the children were recruited from a local school and so would have been more broadly representative of differences in the population than the adult sample (who were all in higher education, and so would be expected to be above average readers, for their age). It stands to reason then that children with lower reading skill would display further reduced rates of lexical processing, given increased difficulties in lexical identification, which would affect their ability to extract information from word *n*+1. As discussed in Section 6.3, if word *n* is slower to be lexically identified, the extraction of information from word *n*+1 is also going to be slower. It is extremely likely, therefore, that variability in reading skill, contingent on lexical processing ability as determined by the quality of lexical representations (e.g., Perfetti, 2007), modulates parafoveal pre-processing. Consistent with this notion, Marx et al. (2016) found that individual reading skill (word-per-minute reading rate) was a greater predictor than school year of a child's ability to extract parafoveal information (i.e., display preview benefit).

Future research, of a longitudinal nature, would be worthwhile in order to truly capture and characterise the developmental change within parafoveal pre-processing of orthography. Such research is needed in order to determine the extent to which reading ability, and the quality of lexical representations, may constrain parafoveal pre-processing.

6.3.2 Implications for Models of Visual Word Recognition

Within Chapter 3 it was found that for both adult and child readers the external letters of a word in preview aid lexical identification processes more so than the internal letters. This finding is consistent with orthographic models of letter position encoding. Adults and children appear, with some flexibility, to code for the most visible letters that best limit the number of competing lexical representation candidates, even for an upcoming word- the external letters (e.g., Spatial Coding model/SOLAR, Davis, 2010a,b; SERIOL, Whitney, 2001). Moreover, with regard to the parafoveal pre-processing of external letters, a highly consistent finding throughout the three experiments (Chapters 3, 4, and 5) was a first-letter bias. This importance of the first letter to lexical identification is entirely consistent, again, with the aforementioned models of letter position encoding; within these models the first letter, due to receiving increased activation/weight, is key to facilitating lexical identification. A limitation of these models, though, is that they do not account for potential developmental change within orthographic processing and letter position encoding.

Grainger and Ziegler's (2011) model of orthographic processing is the only model, at present, that accounts for developmental change within word recognition processes. On the basis of this work, a theoretical framework of orthographic processing within reading has been created (Grainger, Dufau, et al., 2016). Developmental change, over three phases, is explicit within this model and framework: (1) very early reading, involving phonological decoding; (2) parallel letter processing within a word subsequently develops; which leads to the development of (3) two processing routes: coarse-grained (more flexible letter encoding) and fine-grained (limited flexibility to letter encoding and phonology plays a role within this route).

The experiments reported within Chapters 4 and 5 found that processing within the fine-grained route would seem to be comparable between adult and child readers of English. Both groups displayed evidence of phonological recoding (Chapter 4), and it was found that the extraction of phonological information from word *n*+1 was dependent on the extent to which correct orthographic information was present for word *n*+1 (given the fine-grained route's limited flexibility with regard to letter encoding, i.e., word misspellings). The first letter would appear to play a critical role in a reader's (be they beginner or skilled) ability to benefit from phonology in preview. Specifically, the presence of the first letter's correct orthographic code in preview seems to be essential in regard to enabling efficient processing of phonology within the fine-grained route (Chapter 5). This importance of the first letter (first-letter bias) to lexical identification is something which has yet to be incorporated into developmental models of word recognition.

In contrast to the comparable processing within adults' and children's fine-grained routes, there would appear to be some differences within their coarse-grained routes of processing. As mentioned within Section 6.3.1, the results of the experiments reported in Chapters 3 and 5 suggest that 8- to 9-year-old children's parafoveal pre-processing of orthography is slower relative to that of skilled adult readers. This is consistent with the hypothesis of Grainger, Dufau, et al. (2016): becoming a skilled reader involves developing more efficient whole-word identification processes via (pre-)processing of key orthographic features. Whilst such orthographic processing is clear within skilled adult readers, it would seem that beginner readers of English are still developing their coarse-grained routes of processing. In light of the lexical quality hypothesis (e.g., Perfetti, 2007), this is consistent with the suggestion that lexical representations, specifically orthographic lexical representations, are of lower quality within beginner child readers compared to skilled adult readers (e.g., Pagán et al., 2021).

Moreover, this is consistent with research that has shown that there is developmental refinement in orthographic processing (lexical tuning hypothesis; Castles et al., 1999, 2007). The automatic word recognition system (i.e., sight word reading; Ehri, 2005) is required to become more fine-tuned as a reader's lexicon grows. The system can afford, during early word recognition development, to be fairly broadly tuned and accept orthographically similar inputs as candidates for a target word (e.g., *meet* as a candidate for *meat*). This is due to many of these orthographically similar candidate competitors of the word not yet being present in the reader's lexicon (if present at all, i.e., orthographic manipulations typically involve the use of pseudowords, which would not have lexical entries). In this way, the use of less precise criteria (e.g., similar spellings to the correct word are accepted and activate its orthographic representation) allows gains in reading speed without compromising accuracy. As a reader's lexicon grows, typically with age, however, the system must adapt to many more orthographically similar competing words being present in the lexicon (e.g., *meet, seat, meal, neat, mead, etc.* vs. *meat*). The input criterion, therefore, needs to be more precisely tuned in order to cope with this added competition, thereby maintaining maximum accuracy (i.e., spellings are more precisely encoded). As such, lexical representations become, and need to be, of higher quality (e.g., Perfetti, 2007).

Conceivably, the first letter could play a key role in the tuning of orthographic processing. Within adult readers, the orthography of the first letter appears to be precisely encoded, given the significant, immediate costs they displayed to their lexical identification processes when the first letter was substituted in preview (Chapter 5). In contrast, within child readers, the orthography of the first letter would appear to be less precisely encoded, due to differences in the magnitude of the costs associated with the first letter being substituted in preview between the adults' and the children's early processing (Chapter 5). This is consistent with the suggestion made by previous research that the importance of the first letter increases during reading development, driven by increasingly efficient orthographic processing (Grainger, Bertrand, et al., 2016). Consequently, within the word recognition system, the input criterion of the orthography of the first letter would appear to be given increasing priority. As a result, the system becomes increasingly fine-tuned to the orthographic code of the first letter, in order to facilitate lexical identification, due to the informative nature of this letter to word identity (e.g., Clark & O'Regan, 1999; Grainger & Jacobs, 1993).

6.4 Conclusion

Our understanding of parafoveal pre-processing within child readers, compared to adult readers, of English has been advanced by the results of the experiments reported within this thesis. The experimental findings clearly demonstrate that both orthography (Chapters 3 and 4) and phonology (Chapter 4) play important roles in facilitating both adults' and children's lexical identification processes. Furthermore, it was found that orthography determines the extent to which readers are able to benefit from phonological information in preview, with the first letter driving this effect (Chapters 4 and 5). Of note is the consistent finding of the important role the first letter plays in facilitating both adults' and children's ability to lexically identify an upcoming word.

Overall, speed of lexical processing determines not only eye movement behaviour with regard to foveal processing, but also parafoveal pre-processing: the lexical processing of 8- to 9-year-old readers was slower than that of adult readers, resulting in the slower extraction of (orthographic) information from word *n*+1. Nevertheless, typically developing beginner readers of English, even as young as 8-years-old, display a fairly sophisticated ability to extract orthographic and phonological information from word *n*+1, broadly comparable to that of skilled adult readers.

Appendix A

Experimental stimuli

A.1 Stimuli for Experiment 1 (Chapter 3; Milledge et al., 2020)

Experimental sentences and preview conditions (ddd456, 1ddd56, 12ddd6, 123ddd, and dddddd):

The blonde girl spotted the brown *monkey* in the zoo.

(*rackey*, *machey*, *mochiy*, *monhig*, *rachig*)

Tom got an appointment with the nice *doctor* in the hospital.

(*bintor*, *dinfor*, *donfur*, *docfur*, *binfur*)

Peter put clothes in the laundry *basket* ready for washing.

(*hurket*, *burlet*, *barlit*, *baslik*, *hurlik*)

You can find nice fruit in the local *market* on Tuesdays.

(*wonket*, *mondet*, *mandit*, *mardil*, *wondil*)

Kelly always chooses her lucky *number* to play the lottery.

(*savber*, *navter*, *nuvtor*, *numtoc*, *savtoc*)

The man was in grave *danger* as he climbed the mountain.

(*homger*, *domper*, *dampir*, *danpis*, *hompis*)

We saw a large *badger* when we went for a walk last night.

(*hilger*, *bilper*, *balpur*, *badpun*, *hilpun*)

We did not stay much *longer* than you at the birthday party.

(*tumger*, *lumjer*, *lomjar*, *lonjaw*, *tumjaw*)

I like the grey *donkey* that lives in a field behind my house.

(*farkey*, *dartey*, *dortiy*, *dontip*, *fartip*)

Daniel drew a picture with a green *pencil* for his grandma.

(*jumcil*, *pumril*, *pemral*, *penrab*, *jumrab*)

The letter was stuck with a large *magnet* on our fridge door.

(*voynet*, *moyret*, *mayrut*, *magrud*, *voyrud*)

My uncle has a short *temper* and shouts when I’m naughty.

(*dowper*, *towger*, *tewgar*, *temgan*, *dowgan*)

Sue got her hair cut shorter than *normal* and it looked nice.

(*cusmal*, *nusval*, *nosvil*, *norvib*, *cusvib*)

The baby fell asleep after many *tender* kisses from his mum.

(*basder*, *tasfer*, *tesfir*, *tenfim*, *basfim*)

I put lots of silver *tinsel* on the Christmas tree this year.

(*famsel*, *tamrel*, *timrul*, *tinrud*, *famrud*)

The oil was stored in a huge *tanker* until it was needed.

(*lucker*, *tacder*, *tacdor*, *tandos*, *lucdos*)

My football team’s *mascot* is a giant teddy bear in uniform.

(*vixcot*, *mixrot*, *maxret*, *masrel*, *vixrel*)

The little boy is a real *rascal* because he plays jokes on people.

(*wencal*, *renmal*, *ranmul*, *rasmut*, *wenmut*)

My neighbours planted a small *conker* tree in their garden.

(*simker*, *cimber*, *combur*, *conbux*, *simbux*)

The new building has window *ledges* that are painted blue.

(*hubges*, *lubpes*, *lebpas*, *ledpar*, *hubpar*)

Tom cried when his little *finger* got caught in the door.

(*tasger*, *fasyer*, *fisyur*, *finyum*, *tasyum*)

The horse jumped six white *fences* and won the competition.

(*larces*, *farmes*, *fermis*, *fenmix*, *larmix*)

The ambulance took the hurt *victim* quickly to the hospital.

(*surtim*, *vurlim*, *virlom*, *viclon*, *surlon*)

The front *bumper* fell off dad's car today and he was cross.

(*hinper*, *binjer*, *bunjar*, *bumjas*, *hinjas*)

The boss bought a new *dumper* truck for the building project.

(*ticper*, *dicyer*, *ducyar*, *dumyas*, *ticyas*)

The castle has a large *garden* which we like to play in.

(*pocden*, *gochen*, *gachun*, *garhum*, *pochum*)

The space museum had a new model *rocket* ride that was brilliant.

(*wasket*, *rasbet*, *rosbit*, *rocbil*, *wasbil*)

The couple decided to buy a cream *carpet* to go in the bedroom.

(*nimpet*, *cimget*, *camgut*, *cargud*, *nimgud*)

My aunt's chatty *parrot* learns new words very quickly and is very clever.

(*jesrot*, *pescot*, *pascut*, *parcuf*, *jescuf*)

Bob looked down out of the attic *window* to the street below.

(*rasdow*, *waslow*, *wisluw*, *winlum*, *raslum*)

I had some really tasty *turkey* in my sandwich today.

(*dimkey*, *timley*, *tumlay*, *turlag*, *dimlag*)

The children were excited about the great *circus* that was coming to town.

(*mancus*, *canxus*, *cinxes*, *cirxen*, *manxen*)

The photo was of a field with a tiny *piglet* playing in it.

(*qujlet*, *pujdet*, *pijdat*, *pigdab*, *qujdab*)

Alice saw a very prickly *cactus* during her holiday last summer.

(*rintus*, *cinkus*, *cankes*, *cackem*, *rinkem*)

Ben's parents bought a soft *pillow* for his bed last week.

(*gadlow*, *padtow*, *pidtaw*, *piltac*, *gadtac*)

We are looking forward to seeing my clever *sister* come home.

(*romter*, *somler*, *simlur*, *sislun*, *romlun*)

Hannah smiled as the happy *butler* let her into the big house.

(*hadler*, *badfer*, *budfir*, *butfin*, *hadfin*)

He took the empty *carton* from the fridge and threw it away.

(*sixton*, *cixbon*, *caxben*, *carbem*, *sixbem*)

The man ironed his shirt *collar* ready for work the next day.

(*mudlar*, *cudtar*, *codter*, *coltes*, *mudtes*)

Kate peeled and cut the juicy *carrot* ready to put in her dinner.

(*senrot*, *cenmot*, *canmit*, *carmid*, *senmid*)

The lady put a silky *ribbon* onto the dress she was making.

(*makbon*, *raklon*, *riklan*, *riblas*, *maklas*)

The forest was the perfect setting for the family *picnic* last week.

(*yawnic*, *pawric*, *piwrac*, *picrum*, *yawrum*)

Following the instructions, Callum mixed the soft *powder* with a cup of water.

(*junder*, *punber*, *ponbir*, *powbis*, *junbis*)

Mary crawled down the dirty *tunnel* to try to find her football.

(*bacnel*, *tacsel*, *tucsil*, *tunsid*, *bacsid*)

At the animal park there was a huge *walrus* with very long tusks.

(*nibrus*, *wibmus*, *wabmes*, *walmen*, *nibmen*)

The children loved to see the kind *puppet* help his friends.

(*qagpet*, *pagjet*, *pugjot*, *pupjod*, *qagjod*)

The dress was made of a thin *fabric* that was soft to touch.

(*tolric*, *folsic*, *falsuc*, *fabsum*, *tolsum*)

I love to wear my cosy *jumper* for walks when it's cold outside.

(*yawper*, *jawger*, *juwgir*, *jumgis*, *yawgis*)

The man was sent a funny *letter* through the post from his friend.

(*hidter*, *lidber*, *ledbar*, *letban*, *hidban*)

The builders decided to put the strong *ladder* up against the wall.

(*bufder*, *lufter*, *laftir*, *ladtis*, *buftis*)

Jill was proud of the large *turnip* that she had dug up.

(*dacnip*, *tacmip*, *tucmop*, *turmog*, *dacmog*)

I was sent to buy a yellow *pepper* from the supermarket.

(*jagper*, *pagqer*, *pegqur*, *pepqum*, *jagqum*)

It was nearly *winter* and I hoped that it would snow.

(*comter*, *womder*, *wimdar*, *windas*, *comdas*)

Sam looked up at the stars on the clear *summer* night.

(*nicmer*, *sicver*, *sucvar*, *sumvan*, *nicvan*)

A.2 Stimuli for Experiment 3 (Chapter 5; Milledge et al., 2021b)

Experimental sentences and preview conditions (P+ O- and P- O-):

(Note that the sentence frames are grouped here by target word but this was not how they were presented to participants)

We ran in a huge circle round the school field in PE.

The dancers were in a large circle on the stage.

I painted a blue circle on the mug I made for mum.

(*sircle*, *nircle*)

Hannah ate the tasty cereal for breakfast today.

My dad got the full cereal box out of the cupboard.

It is not healthy to eat sugary cereal every day.

(*sereal*, *nereal*)

The bright circus posters were very easy to spot.

Tom heard about the best circus from his friends.

It is exciting when the famous circus comes to town.

(*sircus*, *wircus*)

David didn’t like the mean giant in the film.

They were glad when the young giant helped them.

The happy giant was always eager to make friends.

(*jiant*, *yiant*)

The baby fell asleep after many gentle songs.

My aunt gives me a warm gentle hug whenever I see her.

The lady spoke with a very gentle voice to me.

(*jentle*, *pentle*)

The zookeeper fed the hungry giraffe lots of hay.

The story about the baby giraffe was in the newspapers.

Ben enjoyed seeing the tall giraffe at the zoo.

(*jiraffe*, *piraffe*)

I know that some germs can make you very poorly.

There are bad and good germs inside your tummy.

The teacher's lesson about germs was very interesting.

(*jerms*, *yerms*)

The bus travels between the three cities very slowly.

The bridge between the busy cities was always blocked.

Tim really didn't like the noisy cities at night.

(*sities*, *vities*)

We walked towards the town centre very slowly.

We rode to the city centre on our bikes last night.

Jim helps at an animal rescue centre on weekends.

(*sentre*, *zentre*)

The story was about a brave genie who saved the day.

I jumped when the evil genie appeared out of nowhere.

The magic genie helped us on our way when we got lost.

(*jenie*, *yenie*)

We learned about the last century in history lessons.

I read about the past century in a library book.

The next century should bring exciting new discoveries.

(*sentury*, *xentury*)

The crafty gerbil had managed to escape again.

Sam watched the speedy gerbil run around its cage.

Last night the clever gerbil dug a very long tunnel.

(*jerbil*, *yerbil*)

The girl was a real genius when it came to maths.

Only a true genius could solve the difficult puzzle.

The clear genius of the person was clear to everyone.

(*jenius*, *yenius*)

The city’s small central area was easy to find.

The town’s central square was beautiful in summer.

The book's central character was very popular.

(*sentral*, *mentral*)

The children had many general ideas for the show.

The directions were very general and we got lost.

Lucy asked for some general information about the area.

(*jeneral*, *peneral*)

We became less certain of who would win the prize.

I was quite certain that I knew the right answer.

The man was almost certain he'd made the right choice.

(*sertain*, *mertain*)

The small cinema was always busy at weekends.

Sally went to the quiet cinema with her friends.

They built a new fancy cinema and some shops in town.

(*sinema*, *rinema*)

Appendix B

Supplementary tables, figures, and analyses

B.1 Supplementary Materials for Experiment 1 (Chapter 3; Milledge et al., 2020)

Table B.1.1

Skipping Rates and Standard Deviations (in parentheses) on the Target Word in Each Condition Across All Participants

Group	Condition	Percentage of skips
Adults	123456	6.85% (.25)
	ddd456	1.78% (.13)
	1ddd56	1.97% (.14)
	12ddd6	2.07% (.14)
	123ddd	1.83% (.13)
	dddddd	1.17% (.11)
Children	123456	7.34% (.26)
	ddd456	5.13% (.22)
	1ddd56	4.92% (.22)
	12ddd6	4.55% (.21)
	123ddd	4.52% (.21)
	dddddd	5.49% (.23)

Table B.1.2

Output from Model 1 for First Fixation Duration, Gaze Duration, and Total Reading Time using 123456 as a Baseline for Adults and Children Separately

	First fixation duration				Gaze duration				Total reading time			
	b	SE	*t*	*p*	b	SE	*t*	*p*	b	SE	*t*	*p*
Adults												
123456 (Int)	5.35	.02	227.89	< .001	5.45	.03	213.17	< .001	5.67	.04	142.57	< .001
ddd456	**.13**	**.02**	**5.70**	**< .001**	**.17**	**.03**	**6.74**	**< .001**	**.23**	**.04**	**6.55**	**< .001**
1ddd56	**.13**	**.02**	**5.74**	**< .001**	**.16**	**.03**	**6.39**	**< .001**	**.17**	**.03**	**4.94**	**< .001**
12ddd6	**.10**	**.02**	**4.32**	**< .001**	**.15**	**.03**	**5.90**	**< .001**	**.17**	**.04**	**4.75**	**< .001**
123ddd	**.17**	**.02**	**7.26**	**< .001**	**.21**	**.03**	**8.31**	**< .001**	**.23**	**.04**	**6.41**	**< .001**
dddddd	**.16**	**.02**	**6.88**	**< .001**	**.22**	**.03**	**8.90**	**< .001**	**.26**	**.04**	**7.31**	**< .001**
Children												
123456 (Int)	5.57	.03	163.82	< .001	5.97	.06	92.77	< .001	6.32	.07	90.29	< .001
ddd456	**.08**	**.03**	**2.45**	**.014***	**.11**	**.04**	**2.77**	**.006**	**.15**	**.04**	**3.75**	**< .001**
1ddd56	.02	.03	.47	.639	**.09**	**.04**	**2.14**	**.033***	**.13**	**.04**	**3.26**	**.001**
12ddd6	.01	.03	.19	.847	**.08**	**.04**	**2.07**	**.038***	**.10**	**.04**	**2.48**	**.013***
123ddd	.01	.03	.39	.696	**-.09**	**.04**	**3.23**	**.001**	**.17**	**.04**	**4.44**	**< .001**
dddddd	.06	.03	1.58	.115	-.15	.04	1.82	.069	**.13**	**.04**	**3.33**	**< .001**

Note. The reading time data were log transformed prior to analysis, so the model estimates cannot be directly interpreted. Significant effects are

marked in bold. The syntax, following trimming, for first fixation duration, gaze duration, and total reading time as intercepts only models, for both adults and children, was as follows: *depvar ~ condition + (1|Participant) + (1|targetno)*. The *s denote where the significance levels changed with the use of the *glht* function (i.e., where results went from being significant to non-significant/marginally significant- within first fixation duration: *p* = .059, within gaze duration: *p* = .124 and *p* = .143, respectively, and within total reading time: *p* = .054).

B.2 Supplementary Materials for Experiment 2 (Chapter 4; Milledge et al., 2021a)

Table B.2.1

Single Fixation Probabilities and Standard Deviations (in parentheses) on the Target Word in Each Condition Across All Participants

Group	Condition	Single fixation probability
Adults	Identity	.75 (.50)
	Orthographically similar pseudohomophones	.77 (.49)
	Orthographically similar spelling controls	.76 (.62)
	Orthographically dissimilar pseudohomophones	.74 (.56)
	Orthographically dissimilar spelling controls	.76 (.64)
Children	Identity	.66 (.82)
	Orthographically similar pseudohomophones	.62 (.75)
	Orthographically similar spelling controls	.62 (.77)
	Orthographically dissimilar pseudohomophones	.66 (.86)
	Orthographically dissimilar spelling controls	.68 (.89)

Table B.2.2

Skipping Rates and Standard Deviations (in parentheses) on the Target Word in Each Condition Across All Participants

Group	Condition	Percentage of skips
Adults	Identity	17.33% (.38)
	Orthographically similar pseudohomophones	7.34% (.26)
	Orthographically similar spelling controls	9.30% (.29)
	Orthographically dissimilar pseudohomophones	6.02% (.24)
	Orthographically dissimilar spelling controls	7.43% (.26)
Children	Identity	9.36% (.29)
	Orthographically similar pseudohomophones	9.32% (.29)
	Orthographically similar spelling controls	10.30% (.30)
	Orthographically dissimilar pseudohomophones	8.33% (.28)
	Orthographically dissimilar spelling controls	5.68% (.23)

Table B.2.3

Mean and Standard Deviation (in parentheses) Reading Times on the Orthographically Dissimilar Pseudohomophone Previews

Group	Orthographically dissimilar pseudohomophone previews	First fixation duration (ms)	Single fixation duration (ms)	Gaze duration (ms)
Adults	First letter manipulated	237 (78)	253 (78)	278 (85)
	First letter preserved	235 (76)	244 (77)	258 (80)
Children	First letter manipulated	312 (137)	339 (137)	379 (152)
	First letter preserved	301 (117)	320 (118)	346 (122)

B.3 Supplementary Materials for Experiment 3 (Chapter 5; Milledge et al., 2021b)

Table B.3.1

Single Fixation Probabilities and Standard Deviations (in parentheses) on the Target Word in Each Condition Across All Participants

Group	Condition	Single fixation probability
Adults	Identity	.78 (.85)
	P+ O-	.76 (.79)
	P- O-	.76 (.89)
Children	Identity	.59 (1.52)
	P+ O-	.57 (1.37)
	P- O-	.52 (1.62)

Table B.3.2

Skipping Rates and Standard Deviations (in parentheses) on the Target Word in Each Condition Across All Participants

Group	Condition	Percentage of skips
Adults	Identity	7.91% (.27)
	P+ O-	5.77% (.23)
	P- O-	4.47% (.21)
Children	Identity	4.28% (.20)
	P+ O-	2.53% (.16)
	P- O-	3.52% (.18)

Table B.3.3

Output from Model 1 and the Contrasts for Selective Regression Path Duration and Total Reading Time, using the contr.sdif Function for Presentation

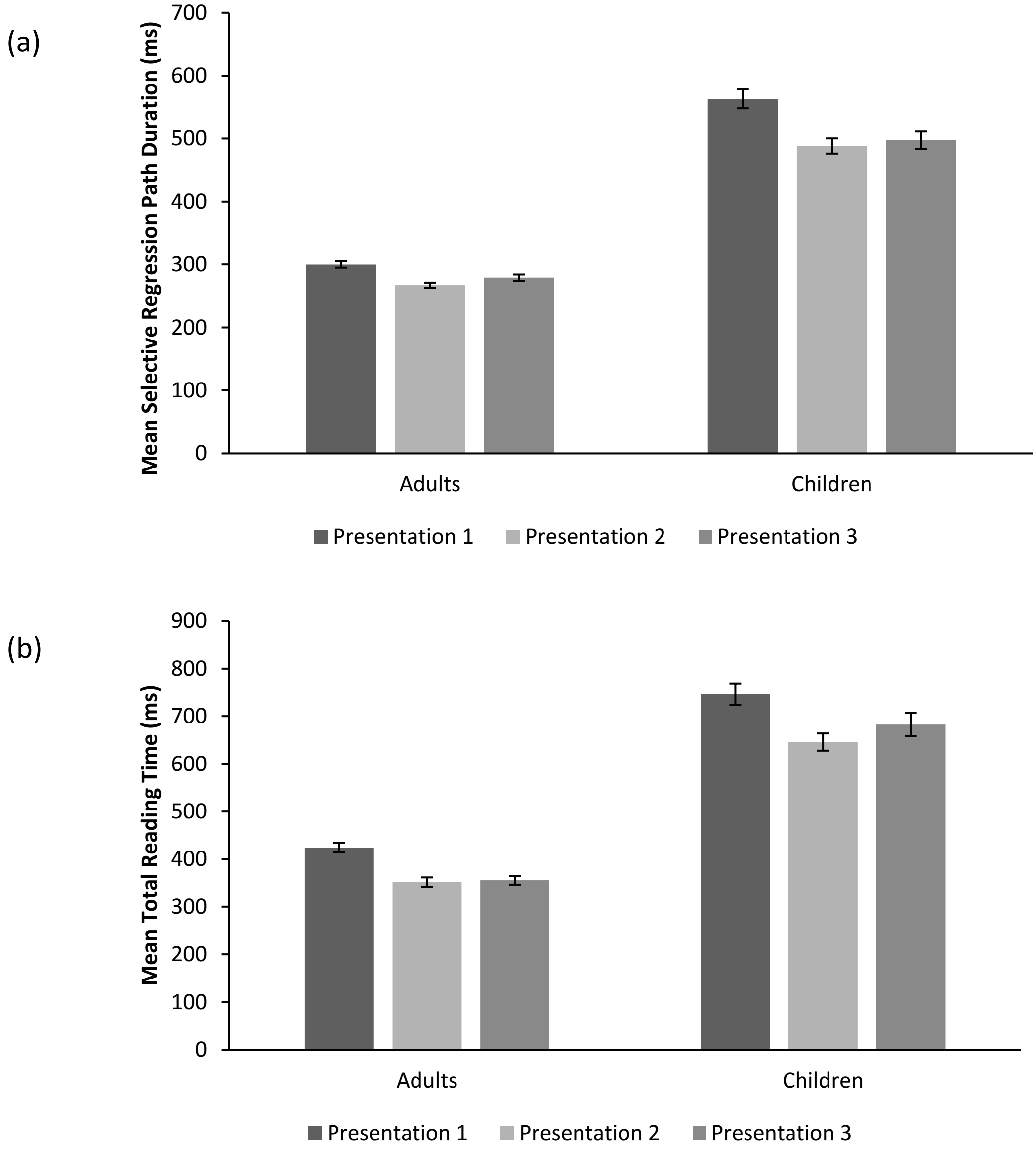
	Selective regression path duration				Total reading time			
	b	SE	*t*	*p*	*b*	*SE*	*t*	*p*
Adults, Identity (Int)	5.47	.05	112.65	< .001	5.69	.06	101.99	< .001
Adults, Children	**.58**	**.06**	**9.03**	**< .001**	**.60**	**.07**	**8.21**	**< .001**
Adults, P+ O-	**.10**	**.03**	**3.88**	**< .001**	**.11**	**.03**	**4.16**	**< .001**
Adults, P- O-	**.16**	**.02**	**6.54**	**< .001**	**.13**	**.03**	**4.82**	**< .001**
Presentation 1-2	**-.11**	**.02**	**-6.35**	**< .001**	**-.16**	**.02**	**-8.21**	**< .001**
Presentation 2-3	.01	.02	.32	.750	.001	.02	.05	.963
Children × P+ O-	-.07	.04	-1.78	.079	**-.11**	**.04**	**-2.75**	**.006**
Children × P- O-	**-.07**	**.03**	**-2.12**	**.037**	-.05	.04	-1.20	.230
Contrasts								
Intercept	5.82	.04	132.55	< .001	6.05	.05	120.28	< .001
Adults, P+ O- vs. P- O-	**-.06**	**.02**	**-2.58**	**.010**	-.02	.03	-.66	.508
Children, P+ O- vs. P- O-	**-.05**	**.02**	**-2.33**	**.020**	**-.08**	**.03**	**-2.94**	**.003**
Presentation 1-2	**-.12**	**.02**	**-7.51**	**< .001**	**-.16**	**.02**	**-8.17**	**< .001**
Presentation 2-3	.01	.02	.62	.537	.0003	.02	.01	.989

Note. The reading time data were log transformed prior to analysis, so the model estimates cannot be directly interpreted. Significant effects are

marked in bold. The *contr.sdif* function (package MASS) was used to set up presentation as a factor. Following trimming, the syntax for selective regression path duration was: *depvar ~ Group * condition + presentation + (1 + condition|Participant) + (1|targetno)*, and the syntax for total reading time, as an intercepts only model, was: *depvar ~ Group * condition + presentation + (1|Participant) + (1|targetno)*. Within the contrasts, after trimming, the syntax for selective regression path duration and total reading time, as intercepts only models, was: *depvar ~ GroupByCond + presentation + (1|Participant) + (1|targetno)*.

Figure B.3.1

Mean Selective Regression Path Durations (a) and Total Reading Times (b) on First, Second, and Third Presentations of Target Words for Both Adults and Children.



References

Adams, M. J., & Huggins, A. W. F. (1985). The growth of children's sight vocabulary: A quick test with educational and theoretical implications. *Reading Research Quarterly, 20*(3), 262-281. https://doi.org/10.2307/748018

Aring, E., Grönlund, M. A., Hellström, A., & Ygge, J. (2007). Visual fixation development in children. *Graefe's Archive for Clinical and Experimental Opthalmology, 245*(11), 1659-1665. https://doi.org/10.1007/s00417-007-0585-6

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*(4), 621-635. doi:10.1017.S0142716403000316

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes, 19*(3), 391–426. https://doi.org/10.1080/01690960344000233

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(2), 416-424. https://doi.org/10.1037/0278-7393.32.2.416

Ashby, J., Yang, J., Evans, K. H. C., & Rayner, K. (2012). Eye movements and the perceptual span in silent and oral reading. *Attention, Perception, & Psychophysics, 74*, 634-640. https://doi.org/10.3758/s13414-012-0277-0

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*(3), 364–390. https://doi.org/10.1016/0010-0285(85)90013-1

References

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. https://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. https://doi.org/10.1016/j.jml.2012.11.001

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

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19*(9), 967-983. https://doi.org/10.1016/0042-6989(79)90222-0

Binder, K. S., Pollatsek, A., & Rayner, K. (1999). Extraction of information to the left of the fixated word in reading. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1162-1172. https://doi.org/10.1037/0096-1523.25.4.1162

Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science, 23*(3), 201-207. https://doi.org/10.1177/0963721414530145

Blythe, H. I., Dickins, J. H., Kennedy, C. R., & Liversedge, S. P. (2018). Phonological processing during silent reading in teenagers who are deaf/hard of hearing: An eye movement investigation. *Developmental Science*, *e12643*, 1-19. https://doi.org/10.1111/desc.12643

Blythe, H. I., Dickins, J. H., Kennedy, C. R., & Liversedge, S. P. (2020). The role of phonology in lexical access in teenagers with a history of dyslexia. *PLOS ONE, 15*(3), e0229934. https://doi.org/10.1371/journal.pone.0229934

Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research, 51*(1), 84-92. https://doi.org/10.1016/j.visres.2010.10.003

Blythe, H. I., & Joseph, H. S. S. L. (2011). Children's eye movements during reading. In S. P. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 643-662). Oxford, UK: Oxford University Press.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular coordination of eye movements during reading in children and adults. *Vision Research, 46*(22), 3898-3908. https://doi.org/10.1016/j.visres.2006.06.006

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research, 49*(12), 1583-1591. https://doi.org/10.1016/j.visres.2009.03.015

Blythe, H. I., Pagán, A., & Dodd, M. (2015). Beyond decoding: Phonological processing during silent reading in beginning readers. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(4), 1244-1252. https://doi.org/10.1037/xlm0000080

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research, 13*(4), 767-782. https://doi.org/10.1016/0042-6989(73)90041-2

Briihl, D., & Inhoff, A. W. (1995). Integrating information across fixations during reading: The use of orthographic bodies and external letters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 55-67. https://doi.org/10.1037/0278-7393.21.1.55

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53-77). Oxford: Oxford University Press.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago: University of Chicago Press.

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology, 97*(3), 165-182. https://doi.org/10.1016/j.jecp.2007.01.006

Castles, A., Davis, C., & Letcher, T. (1999). Neighborhood effects on masked form priming in developing readers. *Language and Cognitive Processes, 14*(2), 201-224. https://doi.org/10.1080/016909699386347

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest, 19*(1), 5-51. https://doi.org/10.1177/1529100618772271

Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology, 59*(3), 209-217. https://doi.org/10.1037/h0087476

Clark, J. J., & O’Regan, J. K. (1999). Word ambiguity and the optimal viewing position in reading. *Vision Research, 39*(4), 843-857. https://doi.org/10.1016/S0042-6989(98)00203-X

Coltheart, V., Laxon, V., Rickard, M., & Elton, C. (1988). Phonological recoding in reading for meaning by adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 387-397. https://doi.org/10.1037/0278-7393.14.3.387

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. doi:10.1037/0033-295X.108.1.204

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65-70. https://doi.org/10.3758/BF03206399

Davis, C. J. (2010a). SOLAR versus SERIOL revisited. *European Journal of Cognitive Psychology, 22*(5), 695-724. https://doi.org/10.1080/09541440903155682

Davis, C. J. (2010b). The spatial coding model of visual word identification. *Psychological Review, 117*(3), 713-758. https://doi.org/10.1037/a0019738

Doctor, E. A., & Coltheart, M. (1980). Children's use of phonological encoding when reading for meaning. *Memory & Cognition, 8*, 195-209. https://doi.org/10.3758/BF03197607

Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 954-969. https://doi.org/10.1037/0096-1523.31.5.954

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

Ehri, L. C. (1998). Word reading by sight and by analogy in beginning readers. In C. Hulme & R. M. Joshi (Eds.), *Reading and spelling: Development and disorders* (pp. 87-111). Mahwah, NJ: Lawrence Erlbaum Associates.

Ehri, L. C. (1999). Phases of development in learning to read words. In J. Oakhill & R. Beard (Eds.), *Reading development and the teaching of reading: A psychological perspective* (pp. 79–108). Oxford, UK: Blackwell.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167-188. https://doi.org/10.1207/s1532799xssr0902_4

Ehri, L. C. (2007). Development of sight word reading: Phases and findings. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 135-154). Oxford, UK: Blackwell.

Ehri, L. C., & Wilce, L. S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology, 75*(1), 3-18. https://doi.org/10.1037/0022-0663.75.1.3

Ehri, L. C., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly, 20*(2), 163-179. https://doi.org/10.2307/747753

Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*(6), 641-655. https://doi.org/10.1016/S0022-5371(81)90220-6

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777-813. https://doi.org/10.1037/0033-295X.112.4.777

Epelboim, J., Booth, J. R., Ashkenazy, R., Taleghani, A., & Steinman, R. M. (1997). Fillers and spaces in text: The importance of word recognition during reading. *Vision Research, 37*(20), 2899-2914. https://doi.org/10.1016/S0042-6989(97)00095-3

Findelsberger, E., Hutzler, F., & Hawelka, S. (2019). Spill the load: Mixed evidence for a foveal load effect, reliable evidence for a spillover effect in eye-movement control during reading. *Attention, Perception, & Psychophysics, 81*, 1442-1453. https://doi.org/10.3758/s13414-019-01689-5

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*(2), 178-210. https://doi.org/10.1016/0010-0285(82)90008-1

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia: Neuropsychological and cognitive studies of phonological reading* (pp. 301-330). Hillsdale, NJ: Erlbaum.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*(1), 71-99. https://doi.org/10.1037/0033-2909.123.1.71

Gagl, B., Hawelka, S., Richlan, F., Schuster, S., Hutzler, F. (2014). Parafoveal preprocessing in reading revisited: Evidence from a novel preview manipulation.

Journal of Experimental Psychology: Learning, Memory, and Cognition, 40(2), 588-595. https://doi.org/10.1037/a0034408

Gough, P. B., & Hillinger, M. L. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society, 30*(1), 179-196. https://doi.org/10.1007/BF02653717

Gough, P. B., Juel, C., & Griffith, P. L. (1992). Reading, spelling, and the orthographic cipher. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 35-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Grainger, J., Bertrand, D., Lété, B., Beyersmann, E., & Ziegler, J. C. (2016). A developmental investigation of the first-letter advantage. *Journal of Experimental Child Psychology, 152*, 161-172. https://doi.org/10.1016/j.jecp.2016.07.016

Grainger, J., Dufau, S., & Ziegler, J. C. (2016). A vision of reading. *Trends in Cognitive Sciences, 20*(3), 171-179. https://doi.org/10.1016/j.tics.2015.12.008

Grainger, J., & Jacobs, A. M. (1993). Masked partial-word priming in visual word recognition: Effects of positional letter frequency. *Journal of Experimental Psychology: Human Perception and Performance, 19*(5), 951-964. https://doi.org/10.1037/0096-1523.19.5.951

Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition, 123*(2), 280-292. https://doi.org/10.1016/j.cognition.2012.01.003

Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology, 2*, 54. https://doi.org/10.3389/fpsyg.2011.00054

Greenberg, D., Ehri, L. C., & Perin, D. (1997). Are word-reading processes the same or different in adult literacy students and third-fifth graders matched for reading level?

Journal of Educational Psychology, 89(2), 262-275. https://doi.org/10.1037/0022-0663.89.2.262

Häikiö, T., Bertram, R., & Hyönä, J. (2010). Development of parafoveal processing within and across words in reading: Evidence from the boundary paradigm. *Quarterly Journal of Experimental Psychology, 63*(10), 1982-1998. https://doi.org/10.1080/17470211003592613

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology, 102*(2), 167-181. https://doi.org/10.1016/j.jecp.2008.04.002

Henderson, J. M., Dixon, P., Petersen, A., Twilley, L. C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 82-97. https://doi.org/10.1037/0096-1523.21.1.82

Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 417-429. https://doi.org/10.1037/0278-7393.16.3.417

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50*(3), 346-363. https://doi.org/10.1002/bimj.200810425

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research, 49*(24), 2948-2959. https://doi.org/10.1016/j.visres.2009.09.012

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(6), 1430-1440. https://doi.org/10.1037/0278-7393.21.6.1430

Inhoff, A. W. (1987). Parafoveal word perception during eye fixations in reading: Effects of visual salience and word structure. In M. Coltheart (Ed.), *Attention and performance* (Vol. 12, pp. 403-420). London: Erlbaum.

Inhoff, A. W. (1989a). Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of Memory and Language, 28*(4), 444-461. https://doi.org/10.1016/0749-596X(89)90021-1

Inhoff, A. W. (1989b). Parafoveal processing of words and saccade computation during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 544–555. https://doi.org/10.1037/0096-1523.15.3.544

Inhoff, A. W., Eiter, B. M., Radach, R., & Juhasz, B. J. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 56*(5), 803-827. https://doi.org/10.1080/02724980244000639

Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 20-34. https://doi.org/10.1037/0096-1523.24.1.20

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*, 431-439. https://doi.org/10.3758/BF03208203

Jared, D., Ashby, J., Agauas, S. J., & Levy, B. A. (2016). Phonological activation of word meanings in Grade 5. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(4), 524-541. https://doi.org/10.1037/xlm0000184

Johnson, R. L., & Dunne, M. D. (2012). Parafoveal processing of transposed-letter words and nonwords: Evidence against parafoveal lexical activation. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 191-212. https://doi.org/10.1037/a0025983

Johnson, R. L., & Eisler, M. E. (2012). The importance of the first and last letter in words during sentence reading. *Acta Psychologica, 141*(3), 336-351. https://doi.org/10.1016/j.actpsy.2012.09.013

Johnson, R. L., Oehrlein, E. C., & Roche, W. L. (2018). Predictability and parafoveal preview effects in the developing reader: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 44*(7), 973-991. https://doi.org/10.1037/xhp0000506

Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 209–229. https://doi.org/10.1037/0096-1523.33.1.209

Johnston, R. S., & Thompson, G. B. (1989). Is dependence on phonological information in children's reading a product of instructional approach? *Journal of Experimental Child Psychology, 48*(1), 131-145. https://doi.org/10.1016/0022-0965(89)90044-1

Jordan, T. R., Almabruk, A. A. A., Gadalla, E. A., McGowan, V. A., White, S. J., Abedipour, L., & Paterson, K. B. (2014). Reading direction and the central perceptual span: Evidence from Arabic and English. *Psychonomic Bulletin & Review, 21*, 505-511. https://doi.org/10.3758/s13423-013-0510-4

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., Gathercole, S. E., & Rayner, K. (2008). Children's and adults' processing of anomaly and implausibility during reading: Evidence from eye movements. *Quarterly Journal of Experimental Psychology, 61*(5), 708-723. https://doi.org/10.1080/17470210701400657

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research, 49*(16), 2078-2086. https://doi.org/10.1016/j.visres.2009.05.015

Joseph, H. S. S. L., Nation, K., Liversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review, 42*(2), 207-222. https://doi.org/10.1080/02796015.2013.12087485

Jouravlev, O., & Jared, D. (2018). Cross-script orthographic and phonological preview benefits. *Quarterly Journal of Experimental Psychology, 71*(1), 11-19. https://doi.org/10.1080/17470218.2016.1226906

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1312-1318. https://doi.org/10.1037/0278-7393.29.6.1312

Juhasz, B. J., & Rayner, K. (2006). The role of age of acquisition and word frequency in reading: Evidence from eye fixation durations. *Visual Cognition, 13*, 846-863. https://doi.org/10.1080/13506280544000075

Juhasz, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Performance, 34*(6), 1560-1579. https://doi.org/10.1037/a0012319

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354. https://doi.org/10.1037/0033-295X.87.4.329

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67–84). Amsterdam: Elsevier Science Publishers.

Kezilas, Y., McKague, M., Kohnen, S., Badcock, N. A., & Castles, A. (2017). Disentangling the developmental trajectories of letter position and letter identity coding using masked priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(2), 250-258. http://dx.doi.org/10.1037/xlm0000293

Kliegl, R., Hohenstein, S., Yan, M., & McDonald, S. A. (2013). How preview space/time translates into preview cost/benefit for fixation durations during reading. *Quarterly Journal of Experimental Psychology, 66*(3), 581-600. https://doi.org /10.1080/17470218.2012.658073

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods, 44*(4), 978-990. https://doi.org/10.3758/s13428-012-0210-4

Kyte, C. S., & Johnson, C. J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology, 93*(2), 166-185. https://doi.org/10.1016/j.jecp.2005.09.003

Leinenger, M. (2014). Phonological coding during reading. *Psychological Bulletin, 140*(6), 1534-1555. https://doi.org/10.1037/a0037830

Lesch, M. F., & Pollatsek, A. (1993). Automatic access of semantic information by phonological codes in visual word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(2), 285-294. https://doi.org/10.1037/0278-7393.19.2.285

Levi, D. M. (2008). Crowding - An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635-654. https://doi.org/10.1016/j.visres.2007.12.009

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., Kentridge, R. W. (2004). Eye movements when reading disappearing text: Is there a gap effect in reading? *Vision Research, 44*(10), 1013-1024. https://doi.org/10.1016/j.visres.2003.12.002

Lukatela, G., & Turvey, M. T. (1994a). Visual lexical access is initially phonological: I. Evidence from associative priming by words, homophones, and pseudohomophones. *Journal of Experimental Psychology: General, 123*(2), 107-128. http://dx.doi.org/10.1037/0096-3445.123.2.107

Lukatela, G., & Turvey, M. T. (1994b). Visual lexical access is initially phonological: 2. Evidence from phonological priming by homophones and pseudohomophones. *Journal of Experimental Psychology: General, 123*(4), 331-353. http://dx.doi.org/10.1037/0096-3445.123.4.331

Luke, S. G., Henderson, J. M., & Ferreira, F. (2015). Children's eye movements during reading reflect the quality of lexical representations: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(6), 1675-1683. https://doi.org/10.1037/xlm0000133

Mancheva, L., Reichle, E. D., Lemaire, B., Valdois, S., Ecalle, J., & Guérin-Dugué, A. (2015). An analysis of reading skill development using E-Z Reader. *Journal of Cognitive Psychology, 27*(5), 657-676. https://doi.org/10.1080/20445911.2015.1024255

Marchbanks, G., & Levin, H. (1965). Cues by which children recognize words. *Journal of Educational Psychology, 56*(2), 57-61. https://doi.org/10.1037/h0021753

Marsh, G., Friedman, M., Welch, V., & Desberg, P. (1981). A cognitive-developmental theory of reading acquisition. In G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (Vol. 3, pp. 199-221). New York: Academic Press.

Marx, C., Hawelka, S., Schuster, S., & Hutzler, F. (2015). An incremental boundary study on parafoveal preprocessing in children reading aloud: Parafoveal masks overestimate the preview benefit. *Journal of Cognitive Psychology, 27*(5), 549-561. https://doi.org/10.1080/20445911.2015.1008494

Marx, C., Hawelka, S., Schuster, S., & Hutzler, F. (2017). Foveal processing difficulty does not affect parafoveal preprocessing in young readers. *Scientific Reports, 7*, 41602. https://doi.org/10.1038/srep41602

Marx, C., Hutzler, F., Schuster, S., & Hawelka, S. (2016). On the development of parafoveal preprocessing: Evidence from the incremental boundary paradigm. *Frontiers in Psychology, 7*, 1-13. https://doi.org/10.3389/fpsyg.2016.00514

Mason, J. M. (1980). When do children begin to read: An exploration of four year old children's letter and word reading competencies. *Reading Research Quarterly, 15*(2), 203-227. https://doi.org/10.2307/747325

Masterson, J., Stuart, M., Dixon, M., Lovejoy, D., & Lovejoy, S. (2003). *The Children's Printed Word Database*. Retrieved from www1.essex.ac.uk/psychology/cpwd

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81*(12), 899-917. https://doi.org/10.1037/h0037368

Mayberry, R.I., del Giudice, A.A., & Lieberman, A.M. (2011). Reading achievement in relation to phonological coding and awareness in deaf readers: A meta-analysis. *Journal of Deaf Studies and Deaf Education, 16*(2)*,* 164-188. https://doi.org/10.1093/deafed/enq049

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*, 578-586. https://doi.org/10.3758/BF03203972

McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics, 25*, 221-224. https://doi.org/10.3758/BF03202990

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. In J. F. Stein (Ed.), *Vision and visual dyslexia* (pp. 251-262). Boston, MA: CRC Press.

Miellet, S., & Sparrow, L. (2004). Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language, 90*, 299-310. https://doi.org/10.1016/S0093-934X(03)00442-5

Milledge, S. V., & Blythe, H. I. (2019). The changing role of phonology in reading development. *Vision, 3*(2), 23. https://doi.org/10.3390/vision3020023

Milledge, S. V., Blythe H. I., & Liversedge, S. P. (2020). Parafoveal pre-processing in children reading English: The importance of external letters. *Psychonomic Bulletin & Review, 28*, 197-208. https://doi.org/10.3758/s13423-020-01806-8

Milledge, S. V., Liversedge, S. P., & Blythe, H. I. (2021a). *Phonological parafoveal pre-processing in children reading English sentences*. Manuscript submitted for publication.

Milledge, S. V., Liversedge, S. P., & Blythe, H. I. (2021b). *The importance of the first letter in children's parafoveal pre-processing in English: Is it phonologically or orthographically driven?* Manuscript submitted for publication.

Morey, R. D., & Rouder, J. N. (2013). BayesFactor: Computation of Bayes factors for common designs (R Package Version 0.9.12-4.2) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/BayesFactor/index.html

Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance, 16*(2), 268-281. https://doi.org/10.1037/0096-1523.16.2.268

Morrison, R. E., & Rayner, K. (1981). Saccade size in reading depends upon character spaces and not visual angle. *Attention, Perception, & Psychophysics, 30*(4), 395-396. https://doi.org/10.3758/BF03206156

O'Regan, J. K. (1979). Saccade size control in reading: Evidence for the linguistic control hypothesis. *Perception & Psychophysics, 25*(6), 501-509. https://doi.org/10.3758/BF03213829

O'Regan, J. K. (1980). The control of saccade size and fixation duration in reading: The limits of linguistic control. *Perception & Psychophysics, 28*(2), 112–117. https://doi.org/10.3758/BF03204335

Pagán, A., Blythe, H. I., & Liversedge, S. P. (2016). Parafoveal preprocessing of word initial trigrams during reading in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(3), 411-432. https://doi.org/10.1037/xlm0000175

Pagán, A., Blythe, H. I., & Liversedge, S. P. (2021). The influence of children's reading ability on initial letter position encoding during a reading-like task. *Journal of Experimental Psychology: Learning, Memory and Cognition.* Advance online publication. https://doi.org/10.1037/xlm0000989

Perea, M., & Acha, J. (2009). Space information is important for reading. *Vision Research, 49*(15), 1994-2000. https://doi.org/10.1016/j.visres.2009.05.009

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on* *resolving lexical ambiguity* (pp. 67-86). Washington, DC: American Psychological Association.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189-213). Amsterdam: John Benjamins.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review, 112*(1), 43-59. https://doi.org/10.1037/0033-295X.112.1.43

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114*(2), 273-315. https://doi.org/10.1037/0033-295X.114.2.273

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14*(1), 174-180. https://doi.org/10.1016/0093-934X(81)90073-0

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 148–162. https://doi.org/10.1037/0096-1523.18.1.148

Pollatsek, A., Perea, M., & Binder, K. S. (1999). The effects of "neighborhood size" in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 1142-1158. https://doi.org/10.1037/0096-1523.25.4.1142

References

Pollatsek, A., & Rayner, K. (1982). Eye movement control in reading: The role of word boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 8*(6), 817-833. https://doi.org/10.1037/0096-1523.8.6.817

Polse, L. R., & Reilly, J. S. (2015). Orthographic and semantic processing in young readers. *Journal of Research in Reading, 38*(1), 47-72. https://doi.org/10.1111/j.1467-9817.2012.01544.x

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from https://www.R-project.org/

Rayner, K. (1975). The perceptual span and peripheral cues during reading. *Cognitive Psychology, 7*(1), 65-81. https://doi.org/10.1016/0010-0285(75)90005-5

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85*(3), 618-660. https://doi.org/10.1037/0033-2909.85.3.618

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*(1), 21-30. https://doi.org/10.1068/p080021

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41*(2), 211-236. https://doi.org/10.1016/0022-0965(86)90037-8

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506. https://doi.org/10.1080/17470210902816461

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology, 40*(4), 473-483. https://doi.org/10.1037/h0080111

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58*, 734-747. https://doi.org/10.3758/BF03213106

Rayner, K., Fischer, M. H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38*(8), 1129-1144. https://doi.org/10.1016/S0042-6989(97)00274-5

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest, 2*(2), 31-74. https://doi.org/10.1111/1529-1006.00004

Rayner, K., & Kaiser, J. S. (1975). Reading mutilated text. *Journal of Educational Psychology, 67*(2), 301-306. https://doi.org/10.1037/h0077015

Rayner, K., Liversedge, S. P., & White, S. J. (2006). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research, 46*(3), 310-323. https://doi.org/10.1016/j.visres.2005.06.018

Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological Science, 14*(4), 385-388. https://doi.org/10.1111/1467-9280.24483

Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology, 12*(2), 206-226. https://doi.org/10.1016/0010-0285(80)90009-2

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading* (2nd ed.). New York, NY: Psychology Press.

Rayner, K., Pollatsek, A., & Binder, K. S. (1998). Phonological codes and eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(2), 476-497. https://doi.org/10.1037/0278-7393.24.2.476

Rayner, K., Sereno, S. C., Lesch, M. F., & Pollatsek, A. (1995). Phonological codes are automatically activated during reading: Evidence from an eye movement priming paradigm. *Psychological Science, 6*(1), 26-32. https://doi.org/10.1111/j.1467-9280.1995.tb00300.x

Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review, 21*, 1067-1072. https://doi.org/10.3758/s13423-014-0582-9

Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review, 17*, 834-839. https://doi.org/10.3758/PBR.17.6.834

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance, 37*(2), 514-528. https://doi.org/10.1037/a0020990

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(6), 1290-1301. https://doi.org/10.1037/0278-7393.30.6.1290

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3*(4), 504-509. https://doi.org/10.3758/BF03214555

Rayner, K., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding wrods with jubmled lettres: There is a cost. *Psychological Science, 17*(3), 192-193. doi:10.1111/j.1467-9280.2006.01684.x

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review, 33*(2), 110-149. https://doi.org/10.1016/j.dr.2013.03.001

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125-157. https://doi.org/10.1037/0033-295X.105.1.125

Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology, 100*(1), 175-191. https://doi.org/10.1037/0022-0663.100.1.175

Rubenstein, H., Lewis, S. S., & Rubenstein, M. A. (1971). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior, 10*(6), 645-657. https://doi.org/10.1016/S0022-5371(71)80071-3

Samuels, S. J., LaBerge, D., & Bremer, C. D. (1978). Units of word recognition: Evidence for developmental changes. *Journal of Verbal Learning and Verbal Behavior, 17*(6), 715-720. https://doi.org/10.1016/S0022-5371(78)90433-4

Schmalz, X., Marinus, E., & Castles, A. (2013). Phonological decoding or direct access? Regularity effects in lexical decisions of Grade 3 and 4 children. *Quarterly Journal* *of Experimental Psychology, 66*(2), 338-346. https://doi.org/10.1080/17470218.2012.711843

Schmalz, X., Marinus, E., Coltheart, M., & Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomic Bulletin & Review, 22*, 1614-1629. https://doi.org/10.3758/s13423-015-0835-2

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics, 74*, 5-35. https://doi.org/10.3758/s13414-011-0219-2

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology, 52*(4), 1021-1046. https://doi.org/10.1080/713755859

Sereno, S. C., & Rayner, K. (1992). Fast priming during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 18*(1), 173-184. https://doi.org/10.1037/0096-1523.18.1.173

Sereno, S. C., & Rayner, K. (2000). Spelling-sound regularity effects on eye fixations in reading. *Perception & Psychophysics, 62*, 402-409. https://doi.org/10.3758/BF03205559

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*(2), 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134*(4), 584-615. https://doi.org/10.1037/0033-2909.134.4.584

Slattery, T. J., Angele, B., & Rayner, K. (2011). Eye movements and display change detection during reading. *Journal of Experimental Psychology: Human Perception and Performance, 37*(6), 1924-1938. https://doi.org/10.1037/a0024322

Snell, J., van Leipsig, S., Grainger, J., & Meeter, M. (2018). OB1-reader: A model of word recognition and eye movements in text reading. *Psychological Review, 125*(6), 969-984. http://dx.doi.org/10.1037/rev0000119

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research, 43*, 219-234. https://doi.org/10.1007/BF00309831

Sperlich, A., Schad, D. J., & Laubrock, J. (2015). When preview information starts to matter: Development of the perceptual span in German beginning readers. *Journal of Cognitive Psychology, 27*(5), 511-530. https://doi.org/10.1080/20445911.2014.993990

Spragins, A. B., Lefton, L. A., & Fisher, D. F. (1976). Eye movements while reading and searching spatially transformed text: A developmental examination. *Memory & Cognition, 4*, 36-42. https://doi.org/10.3758/BF03213252

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: some current controversies. *Trends in Cognitive Sciences, 5*(4), 156-163. https://doi.org/10.1016/S1364-6613(00)01619-3

Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition, 30*(2), 139-181. https://doi.org/10.1016/0010-0277(88)90038-8

Taylor, S. E. (1965). Eye movements while reading: Facts and fallacies. *American Educational Research Journal, 2*(4), 187-202. https://doi.org/10.3102/00028312002004187

Tiffin-Richards, S. P., & Schroeder, S. (2015a). Children's and adults' parafoveal processes in German: Phonological and orthographic effects. *Journal of Cognitive Psychology, 27*(5), 531-548. https://doi.org/10.1080/20445911.2014.999076

Tiffin-Richards, S. P., & Schroeder, S. (2015b). Word length and frequency effects on children's eye movements during silent reading. *Vision Research, 113*, 33-43. https://doi.org/10.1016/j.visres.2015.05.008

Valle, A., Binder, K. S., Walsh, C. B., Nemier, C., & Bangs, K. E. (2013). Eye movements, prosody, and word frequency among average- and high-skilled second-grade readers. *School Psychology Review, 42*(2), 171-190. https://doi.org/10.1080/02796015.2013.12087483

Van Orden, G. C. (1987). A rows is a rose: Spelling, sound, and reading. *Memory & Cognition, 15*, 181-198. https://doi.org/10.3758/BF03197716

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N + 1 and word N + 2 during reading: A critical review and Bayesian meta-analysis. *Psychonomic Bulletin & Review, 24*, 666-689. https://doi.org/10.3758/s13423016-1147-x

Vasilev, M. R., Slattery, T. J., Kirkby, J. A., & Angele, B. (2018). What are the costs of degraded parafoveal previews during silent reading? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(3), 371-386. https://doi.org/10.1037/xlm0000433

Vasilev, M. R., Yates, M., & Slattery, T. J. (2019). Do readers integrate phonological codes across saccades? A Bayesian meta-analysis and a survey of the unpublished literature. *Journal of Cognition, 2*(1), 43. https://doi.org/10.5334/joc.87

Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *Quarterly Journal of Experimental Psychology, 67*(4), 703-727. https://doi.org/10.1080/17470218.2013.826258

Veldre, A., & Andrews, S. (2018). How does foveal processing difficulty affect parafoveal processing during reading? *Journal of Memory and Language, 103*, 74-90. https://doi.org/10.1016/j.jml.2018.08.001

Vellutino, F.R., Fletcher, J.M., Snowling, M.J., & Scanlon, D.M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*(1), 2-40. https://doi.org/10.1046/j.0021-9630.2003.00305.x

Vitu, F., & McConkie, G. W. (2000). Regressive saccades and word perception in adult reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 301-326). Amsterdam: Elsevier.

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading: Eyemovement characteristics in scanning letter strings and reading texts. *Perception & Psychophysics, 57*, 352-364. https://doi.org/10.3758/BF03213060

Wang, A., Yan, M., Wang, B., Jia, G., & Inhoff, A. W. (2020). The perceptual span in Tibetan reading. *Psychological Research*. https://doi.org/10.1007/s00426-020-01313-4

Wechsler, D. (2005). *Wechsler Individual Achievement Test: Second UK Edition (WIAT-II UK)*. London, UK: Pearson.

White, S. J. (2008). Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 34*(1), 205-223. https://doi.org/10.1037/0096-1523.34.1.205

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word-beginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1261-1276. https://doi.org/10.1037/0096-1523.34.5.1261

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin & Review, 12*, 891-896. https://doi.org/10.3758/BF03196782

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review, 8*, 221-243. https://doi.org/10.3758/BF03196158

Williams, J. P., Blumberg, E. L., & Williams, D. V. (1970). Cues used in visual word recognition. *Journal of Educational Psychology, 61*(4), 310-315. https://doi.org/10.1037/h0029570

World Literacy Foundation. (2015). *The economic & social cost of illiteracy: A snapshot of illiteracy in a global context.* Retrieved from https://worldliteracyfoundation.org/wp-content/uploads/2015/02/WLF-FINAL-ECONOMIC-REPORT.pdf

Yan, M. Li, H., Su, Y., Cao, Y., & Pan, J. (2020). The perceptual span and individual differences among Chinese children. *Scientific Studies of Reading, 24*(6), 520-530. https://doi.org/10.1080/10888438.2020.1713789

Yan, G., Liu, M., Zhu, M., Sainan, L., Yongsheng, W., & Liversedge, S. P. (2019). Orthographic, phonological and semantic preview benefit for Chinese children aged 7-8, 8-9, 9-10 and 10-11 years. *Journal of Eye Movement Research, 12*(7), 185.

Yan, M., Richter, E., Shu, H., & Kliegl, R. (2009). Chinese readers extract semantic information from parafoveal words during reading. *Psychonomic Bulletin & Review, 16*, 561-566. https://doi.org/10.3758/PBR.16.3.561

Yan, M., Wang, A., Song, H., & Kliegl, R. (2019). Parafoveal processing of phonology and semantics during the reading of Korean sentences. *Cognition, 193*, 104009. https://doi.org/10.1016/j.cognition.2019.104009

Yang, J., Wang, S., Tong, X., & Rayner, K. (2012). Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading. *Reading and Writing, 25*, 1031-1052. https://doi.org/10.1007/s11145-010-9281-8

Zhang, M., Liversedge, S. P., Bai, X., Yan, G., & Zang, C. (2019). The influence of foveal lexical processing load on parafoveal preview and saccadic targeting during Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance, 45*(6), 812-825. http://dx.doi.org/10.1037/xhp0000644

Zhou, W., Shu, H., Miller, K., & Yan, M. (2018). Reliance on orthography and phonology in reading of Chinese: A developmental study. *Journal of Research in Reading, 41*(2), 370-391. https://doi.org/10.1111/1467-9817.12111

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., Saine, N., Lyytinen, H., Vaessen, A., & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science, 21*(4), 551-559. https://doi.org/10.1177/0956797610363406

Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society of London B, Biological Sciences, 369*(1634), 20120397. https://doi.org/10.1098/rstb.2012.0397