Investigating Transposed Word Effects in Reading

by

Petar Filipov Atanasov

A thesis submitted in partial fulfilment for the requirements for the degree of Doctor of Philosophy at the University of Central Lancashire

January 2025

RESEARCH STUDENT DECLARATION FORM

Type of Award: Doctor of Philosophy

School of Psychology and Humanities

1. Concurrent registration for two or more academic awards

I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution.

2. Material submitted for another award

I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work.

3. Collaboration

Where a candidate's research programme is part of a collaborative project, the thesis must indicate in addition clearly the candidate's individual contribution and the extent of the collaboration. Please state below:

4. Use of a Proof-reader

No proof-reading service was used in the compilation of this thesis.

Signature of Candidate Petar Atanasov

Print name: Petar Filipov Atanasov

Acknowledgments

There are many people without whose support, guidance, patience, and faith in me, this thesis would not be possible, and I would not have the privilege to be a postgraduate student to begin with. I want to start chronologically by expressing my immense gratitude to my family, especially my parents, Stefka and Filip, who have never stopped believing in me since I first drew breath approximately twenty-six years ago. My parents have always put me and my needs first for as long as I can remember, their unwavering support and unconditional faith not only in my abilities and potential but also in my intentions and vision have consistently given me confidence and allowed me to persevere regardless of the size and shape of the proverbial stones life has thrown at me. Thank you, Mom and Dad, this one is for you!

Next, I want to thank the many friends, classmates, and teachers who have helped me develop as an individual both at school and at the Aikido dojo where I trained for the better part of a decade during my teenage years. There are too many names, too many stories, and certainly too many fond memories to share here, but I am confident that I would not have become the person I am today without you.

Next, turning to my undergraduate degree days, again, there are too many people who helped me along the way to name you all. I am grateful to every single one of you. I do want to specifically thank Dr. Joshua Ng. whom I met during the first year of my undergraduate degree. You helped me to learn how to deal with the immense pressure I was experiencing from the transition to life in another country, away from home, and use it to grow both as a person and as an aspiring academic.

Now I would like to turn to my supervisors, Professor Simon P. Liversedge and Dr. Federica Degno. The past three and a half years have been a rollercoaster of events and I cannot begin to thank you enough for how much you have taught me, inspired me, and enabled me to develop as a researcher. You have poured your time, efforts, energy, and expertise into this project and me and I am grateful for everything you have done for me, doubly so knowing how extremely busy you are and have been. Regardless of the circumstances, I have always been able to come to you for advice and support. Thank you, it has been a privilege to learn from you.

Next, I would like to thank all of the other Ph.D. students in the Perception Cognition and Neuroscience research group. All of you are amazing. You have been there with me to share the joy of a program finally doing what it should, the frustration that comes with a line of code not working for days, and realising the whole issue lies in a missing comma, and so many memorable and less memorable everyday situations we have to deal with. Specifically, thanks to Dr. Zoe Hughes, Laura Rettie, and Rohanna Sells for being the best office mates, colleagues, and friends I could hope for. For all my other postgraduate colleagues, regardless of how long I have known you, thank you for your support as well.

I would like to also thank Dr. Ying Fu and Dr. Mengsi Wang, without whom I would never have been able to finish collecting the data for my last two studies and I would likely still be in the lab now. I am grateful not only for your help but also for your friendship and support. On this note, thanks also to all of my participants without whom I would not be submitting this thesis at all.

I need to also express my gratitude to my girlfriend's family for always welcoming me with open arms and making it feel like I have a home away from home. Ian and Jacqui – Thank you!

Finally, I want to say a very big thank you to my girlfriend Millie. Your understanding, compassion, love, and in no small part, honesty, especially when it comes to how I always forget to hoover, have helped me beyond what I can express with words. You are honestly, the best!

Table of Contents

List of tables

List of figures

Abstract

According to the recently established OB1 Reader model (Snell, van Leipsig, et al., 2018), word position coding is noisy, and words are assigned to sentence positions via a spatiotopic sentence-level mechanism on the basis of length cues and syntactic expectations. A critical prediction that follows from this assumption is that readers should sometimes fail to detect an ungrammaticality created by transposing two adjacent words within a sentence. In line with this prediction, Mirault et al. (2018) initially found that readers are worse at judging the (un)grammaticality of sentences containing a word transposition only, versus sentences containing a word transposition and a syntactically illegal final word – a *Transposed-Word* effect. Since then, word transpositions have become increasingly relevant to the investigation of whether lexical processing is serial or parallel. The present thesis was aimed at investigating how word transpositions may influence moment-to-moment online cognitive processing during silent sentence reading via eye-tracking. The first experiment explored the influence of word transpositions and final word ungrammaticality on eye movements and grammaticality decisions. The second experiment explored the effects of word transpositions in the parafovea versus the fovea on eye movements and grammaticality decisions via the boundary paradigm (Rayner, 1975). The third experiment investigated the influence of word transpositions in comparison to letter masks in the parafovea on eye movements via the boundary paradigm as well. Across the three experiments, readers exhibited no sensitivity to the presence of a word transposition in parafoveal vision. However, from Experiment 1 and Experiment 2, it was clearly evident that upon fixating the first of the two transposed target words in a sentence, readers experienced significant and immediate disruption to processing. Furthermore, findings from Experiment 2 and Experiment 3 indicated that readers were sensitive to a preview change created by changing the order of the two target words in parafoveal versus foveal vision. Importantly, that sensitivity was likely due to the visual and orthographic rather than the syntactic mismatch between preview and targets and was driven primarily by the invalid preview of the first rather than the second target word. Overall, the results from the three experiments suggest that words are lexically processed serially and sequentially rather than in parallel. Moreover, readers do indeed sometimes fail to detect word transpositions, however, that is likely not driven by the processing mechanisms proposed by the OB1 Reader model. Consequently, current serial attention shift computational models of reading need to be expanded and modified in order to account for why readers fail to detect word transpositions under serial lexical processing assumptions.

Chapter 1. Introduction

1.1. Reading as an essential cognitive process for everyday life

The ability to read has been consistently linked with a range of academic (Kirsch et al., 2003), professional, and health (Dugdale & Clark, 2008; Parsons & Bynner, 2008) outcomes in modern society. For instance, in the United Kingdom, children with lower reading ability have consistently scored lower on GCSE examinations compared to their peers who are better at reading (DfE, 2015). Beyond childhood, reading ability in adulthood has been positively associated with employment and well-being in that better readers are more likely to earn above-average salaries and be employed in skilled jobs while also being more happy and likely to use preventative healthcare than adults with low reading ability (e.g., Martinez & Fernandez, 2010; World Literacy Foundation, 2018). In short, reading is a very important skill that is critical to, and positively associated with, multiple aspects of daily life.

Despite it being so important, silent reading has only existed for approximately the past two thousand years with some of the earliest surviving mentions of reading without vocalising being from the series of books *Confessions* by St Augustine which were written between 397-398 AD (see Saenger, 1997). Furthermore, it has only been in the last century or so, that literacy (the ability to read and write) has rapidly increased from around 20% of the population in the world in the early 1900s to over 85% of the world population in the 2000s for adults aged over 15 years (van Zanden et al., 2014). In parallel with this increase in literacy, cognitive scientists have become interested in understanding the psychological mechanisms and systems by which reading occurs with early scientific works dating as far back as Huey's *Psychology and Pedagogy of Reading* from 1908.

Since the 1970s, the eye-tracking methodology has become central to reading research as eye movements provide a fine-grain, online, ecologically valid way to measure reading that allows researchers to gain an insight into how readers process written text (Rayner, 1998, 2009). Besides empirical studies, multiple oculomotor control models of reading have been developed to provide mechanistic accounts of how both visual and linguistic characteristics of text influence reading as reflected by both temporal (e.g., how long readers look at a word or words) and spatial (e.g., which words are directly fixated, where they are fixated and how far the eyes move forward, or backward, in a text) aspects of eye movements.

In the following section, I will provide a brief overview of the physiology of the human eye. Subsequently, I will discuss the utility of the eye movement method in investigating the cognitive processes that underpin reading. I will then provide a brief overview of the language processes that underlie reading before discussing several computational models that aim to predict how visual and linguistic characteristics of text affect eye movement control via processing mechanisms to determine when and where the eyes move through text. Afterward, I will provide an overview of the recently established *Transposed-Word* effect (e.g., Mirault et al., 2018) which is central to the three experiments of my Ph.D., and I will establish the rationale and motivation for each of the studies I have conducted.

1.2. Physiology of the human eye and eye movement control

In this section, I will briefly consider the structure of the eye with a focus on how and why the eye moves when exposed to visual stimuli. This is a deliberate choice as understanding the structure of the human eye, how it extracts information from the surrounding environment in general, and particularly during reading, and how that information is transmitted to the brain is crucial to allow for understanding of subsequent sections.

The eye is amongst the most complex organs in the human body. It is a spherical organ with a small aperture, the pupil, that permits light to enter the sphere. Light passes from the environment into the eye and hits the inner, light-sensitive surface (the retina). The light-sensitive surface functions to convert the light signal into a neural signal that is then delivered to the human brain to allow for further visual and cognitive processing. As light feeds through the pupil to the back of the eye, it passes through the cornea, a transparent layer of tissue at the front of the eye that refracts light. The light then moves through the lens which focuses the light onto the retina (e.g., Maurice, 1984; Muller et al., 2003). The iris is a structure that controls the size of the pupil by dilating or constricting it, thereby controlling the amount of light that falls on the retina. When light reaches the retina, it is transferred to regions of the brain where visual processing takes place, and the encoding of information from the environment can then begin.

Next, let us focus on the retina, the area at the back of the inside of the eye, where light from the environment falls after passing through the pupil. Such light is detected by rod and cone cells which are stimulated. This represents the start of the process by which visual stimuli are encoded and represented. Estimates regarding the number of cones and rods vary across the literature ranging from 4.6 million cones and 97 million rods (Curcio et al., 1990) to 7 million cones and 120 million rods on average (Molday & Moritz, 2015). Rod cells are important for our perception of light and function best in dim light conditions. Cone cells, on the other hand, are more sensitive to oscillations in light intensity and as such function best in conditions with more ambient light. Furthermore, the concentration of cones in the retina is not uniform (Chui et al., 2008; Song et al., 2011), and the lower the concentration of cone cells, the lower the visual acuity (detailed visual sensitivity) in the corresponding area of the retina.

The variations in rod and cone cell concentration across the retina are inversely proportional, and according to this variability, the retina may be considered to be comprised of three loosely defined zones of visual acuity: the fovea – the area that overlaps almost fully with the macula (the central area of the retina) with a diameter of approximately 2 degrees of visual angle. This area has the highest proportion of cone cells in the retina and therefore has the highest visual acuity; the parafovea – the area surrounding the fovea between roughly 2 and 5 degrees of visual angle with visual acuity proportionally and rapidly decreasing with increasing distance from the centre of the macula (see Balota & Rayner, 2012; Rayner, 1998); the third area of the retina is termed the periphery, the rest of the retina, which has the lowest visual acuity due to its lowest concentration of cone cells (see Rosenholtz, 2016 for a review).

Importantly, during reading of most languages, text is horizontally extended across the page or screen, and it is static. In almost all orthographies, words appear in sequences (usually horizontally, though sometimes vertically) and light from the words falls over the extent of the retina (i.e., where there is high visual acuity at the fovea, poorer acuity in the parafovea and poorest acuity in the periphery). This means that to process words efficiently, the eye(s) must be repeatedly re-oriented, very precisely, such that light from successive words in a text is caused to fall on the fovea (Rayner, 2009). To achieve this, humans make a series of fixations and saccades when they read.

Saccades are rapid rotations of the eyeball whose role under most circumstances is to bring the next chunk of text (e.g., the next word or so) into foveal vision so that it may be processed efficiently under optimal viewing conditions (Liversedge & Findlay, 2000; Rayner, 2009). Saccades are targeted movements and can be thought of as *ballistic* (e.g., Liversedge & Findlay, 2000). They are also very fast movements that can reach velocities of up to 600º per second (e.g., Leigh & Zee, 2006). For example, a saccade that moves the eyes 5º to the right would take approximately 40-50ms to complete (e.g., Rayner, 1978; Rayner, 1998). Additionally, progressive saccades, during reading (not considering saccades between lines of text – i.e., return sweeps) have an average length of 2º (Rayner, 2009), or seven to nine character spaces (e.g., Morrison & Rayner, 1981) regardless of the distance at which text is viewed. Consequently, saccadic targets are usually selected from the upcoming words in the parafovea. Importantly, the majority of saccades follow the direction of reading, while a minority (e.g., 10-15%) of saccades are regressions to the immediately preceding word, or even further back into the text (see Rayner, 2009; Schotter et al., 2012).

Between saccades, there are periods of (relatively) stable gaze, called fixations, during which, detailed visual and linguistic information is extracted and processed. Readers typically make four fixations on average each second, with fixation durations varying from approximately 60 to 500ms (Rayner, 2009; Rayner et al., 2012). While some movement occurs during fixations, such as tremors, drifts, and microsaccades (e.g., see Martinez-Conde et al., 2004 for a review), the role of these fixational movements in cognitive processes is controversial and likely not primarily indicative of moment-to-moment ongoing cognitive processing. It is during fixations that new visual and linguistic information can be encoded, as light from the fixated word falls onto the fovea consistently for an extended period of time. Several decades of research have clearly shown that the durations of fixations, saccade length, and fixation probabilities are tightly linked with ongoing cognitive processing such that, for instance, the more difficult a word is to process, the longer readers fixate on it, the more likely they are to fixate it multiple times, and the less likely they are to skip it (Liversedge & Findlay, 2000; Rayner, 1998, 2009).

In addition, studies have shown that it takes at least approximately 125-175 milliseconds to prepare to launch a saccade (e.g., Becker & Jürgens, 1979; Rayner, 1998) meaning that a saccadic program is prepared simultaneously with the processing of the fixated word. That is to say, multiple moment-to-moment cognitive processes unfold during each fixation. The combination of saccades and fixations allows readers to efficiently process text. This is because saccades reorient the eye such that light from the word that readers aim to process falls onto the fovea while fixations are the periods during which, information pertaining to the foveated word and (to an extent) the word or words in the parafovea is processed and the next saccade is programmed.

Research has shown that little-to-no new information is extracted during saccades (e.g., Bremmer et al., 2009; Matin, 1974; Wurtz, 2008). This is because a phenomenon known as saccadic suppression occurs during a saccade, whereby visual input due to light falling onto the retina is suppressed. The precise mechanisms of saccadic suppression are still debated and are beyond the scope of the current thesis and they will not be discussed further. While no new information is encoded during a saccade, several studies have shown that ongoing cognitive processing based on already encoded information during previous fixations does continue during a saccade under most conditions (e.g., Irwin, 1998; Irwin et al., 1995; Yatabe et al., 2009).

Multiple animal studies have established that there are two distinct neural systems that govern saccade generation (e.g., Van Gisbergen et al., 1981; see Wurtz & Goldberg, 1989 for a review; Findlay & Walker, 1999 for a discussion) with one system controlling when a saccade is launched (a *when* decision) coinciding with the duration of fixations, and a separate system controlling where the saccade is aimed to land (a *where* decision). Saccadic generation is directly controlled by several populations of neurons in the brain stem. Specifically, omnipause cells in the caudal pons fire at high rates during periods of stable gaze but are inhibited for the duration of the saccade and the preceding 10 to 12 milliseconds (e.g., Butter-Ennever et al., 1988; Horn et al., 1994; Yoshida et al., 1999). In addition, burst cells in the reticular formation of the brain stem directly drive the saccadic movements by transmitting excitatory or inhibitory signals to the motoneurons that trigger one of the three pairs of muscles in the eyes to generate the movement (e.g., Scudder et al., 1988; Strassman et al., 1986a, b).

Besides the brain stem, the superior colliculus (e.g., Munoz & Wurtz 1992; 1993a; 1993b; Munoz 2002; Wurtz 1996), the frontal eye fields (e.g., Segraves 1992; Hanes & Schall, 1996) and several other populations of neurons in other brain structures (e.g., parietal, and frontal cortex: Goldberg & Segraves 1989; Hyvarinen 1982; Sakata et al., 1980) have also been linked to saccade generation. Importantly studies have shown that similarly to the brain stem, there are distinct populations of neurons in the superior colliculus which are involved in the preparation, generation, and targeting of saccades respectively (e.g., Basso & Wurtz 1998; Horwitz and Newsome 2001; Munoz & Wurtz 1995; see Wurtz, 2000 for a review). All of these studies provide strong support for the notion that the timing and targeting of saccades are the product of two separate, but interactive neurophysiological systems via specialist and distinct neuron populations in several brain areas that direct saccade timing and saccade targeting (e.g., Findlay & Walker, 1999). Consequently, in eye movement research investigating reading, it is important to consider what both fixation durations and saccade characteristics such as length and direction might reveal with respect to the complex cognitive processes that occur as text is read.

Eye movements are an ecologically valid way to study reading because they occur naturally as humans read text and as mentioned previously, are known to index online cognitive processing (e.g., Rayner, 2009). In addition, eye movements can be recorded with high precision and noninvasively via modern video-based eye trackers (e.g., Eyelink 1000Plus – SR Research). Such machines allow researchers to measure both fixation and saccade characteristics at a high spatial (i.e., 0.01°) and temporal (i.e., one sample per millisecond) resolution. In other words, not only are eye movements an ecologically valid and fine-grained way to measure the online cognitive processes during reading, but because of technological capabilities, they can be measured very accurately and fairly easily. These two factors have made eye-tracking a preferred method of many cognitive scientists for studying reading under ecologically valid conditions for the past fifty years (Rayner, 1978, 1998, 2009). In the following paragraphs, I will outline (some of) the eye movement measures that have been used to investigate the factors that influence *when*, and *where*, the eyes move through text. Subsequently, I will focus on what cognitive factors influence the *when* and *where* oculomotor decisions in reading as reflected in the patterns associated with fixations and saccades.

1.3. Eye movement research in reading

Eye-tracking research investigating reading, over the past fifty years, has utilised a range of standardised measures (see Rayner, 2009) reflecting both the time it takes to process a word or phrase, the size and direction of eye movements, and the probability of making a fixation. Eye movement measures can be considered both at the local level of individual words as well as at a global level (e.g., across all words in a sentence or paragraph of text). For example, multiple studies have found that longer words receive more than one fixation more often than shorter words (e.g., Just & Carpenter, 1980), while more frequently occurring words, as estimated by corpus studies (e.g., Balota et al., 2007; Francis & Kućera, 1982; van Heuven et al., 2014), are fixated for less time than less frequently occurring words (e.g., Henderson & Ferreira, 1990, 1993; Hyönä & Olson, 1995; Inhoff & Rayner, 1986) which has been taken as evidence that longer and less frequent words are harder to process. In other words, the underlying assumption when using eye movement measures is that they provide a temporally precise reflection of the difficulty (or ease) readers experience as they process written text. To that end, early local reading time measures such as first fixation duration (FFD: the duration of the first fixation on a word before the eyes have moved to subsequent words), single fixation duration (SFD: the duration of the fixation when only one fixation was made on a word during first-pass reading), and gaze duration (GD: the sum of all fixations on a word during first-pass reading before the eyes move away to another region), as well as fixation probability measures such as refixation probability (RP: the probability to fixate a word more than once before moving the eyes to another word during first-pass reading), are thought to reflect how difficult a word is to process. In contrast, skipping probability (SP: the probability of skipping a word during first-pass reading) has been associated with the tendency to not directly fixate short words, particularly function words such as *the* (see Drieghe, 2008; Drieghe et al., 2004, 2005, 2019), highly predictable words given previous context (e.g., Balota et al., 1985; Drieghe et al., 2005; Ehrlich & Rayner, 1981; Rayner & Well, 1996; Schustack et al., 1987; Vitu, 1991) or highly frequent words (e.g., Kliegl et al., 2004; Rayner & Duffy, 1986; Schilling et al., 1998) suggesting that such words can be lexically identified without being directly fixated. It is worth noting as well, that all of these measures can be computed not just for first-pass reading but also for any subsequent reading cycles (e.g., if a portion of text is read three separate times, then estimates regarding the reading times for each of the three separate occasions can be calculated). However, measures on second or even third-pass reading would not be indicative of early lexical processing and instead may reflect later higher-order processes.

Later measures such as go-past time (GPT: the duration of all fixations on the word and any regressive fixations on words to the left of the current word, before the eyes move to the right), total viewing time (TVT: the total time spent reading the word), the probability to regress into a word (RIn: the probability of fixating a word after having fixated a word to the right), and the probability to regress out of a word (ROut: the probability that the eyes move from the currently fixated word to another word to the left during first-pass reading) have, in contrast, often been shown to reflect the effort associated with integrating an identified word into context, (e.g., Ehrlich & Rayner, 1981; Hyona, 1993; Inhoff, 1984; Rayner & Well, 1996) or recovering from an initial erroneous interpretation of preceding context (e.g., Frazier & Rayner, 1982; Rayner & Frazier, 1987; Rayner et al., 1992). In addition, as noted earlier, most of these measures can be computed over a larger portion of text (e.g., a phrase comprising several words) or even the whole sentence to yield an estimate of how difficult that portion of text (or sentence) is to process overall. In short, multiple eye movement measures are widely used to investigate different aspects of reading from lexical identification to sentential integration with the basic premise that longer reading times and more fixations can be interpreted as reflecting processing difficulty.

Figure 1.1 below shows a hypothetical pattern of eye movements during sentence reading from a fictitious participant. From that pattern, most of the aforementioned local eye movement measures such as first fixation duration or gaze duration can be computed for each word. For example, the reader made their first and only fixation on the first word of the sentence - the preposition *During*. Hence, for this word, FFD coincides with SFD and GD, that correspond to 225ms. The reader then skipped the determiner *the* via a saccade that was 12 characters long, landing on the noun *journey*. When calculating skipping probability (e.g., of the second word in the hypothetical sentence), the computation is based on the proportion of trials on which a word was skipped and is most often reported as a percentage value. In the example, the following fixations from 3 to 7 were all progressive in the direction of reading. Note that fixations 5 and 6 landed on the same word. Therefore, fixation 6 was a refixation in the first-pass reading of the verb *impressed*. Consequently, the FFD for this word was equal to the duration of fixation 5 (i.e., 160ms), since the word received two fixations, SFD was not computed because there was more than one first-pass fixation on the word. However, GD was computed as the sum of fixations 5 and 6 (i.e., 160ms + 215ms = 375ms). Additionally, this hypothetical trial would contribute to the calculation of refixation probability on the word *impressed* in that sentence. Similarly to skipping probability, the probability to refixate would be calculated based on the proportion of trials on which *impressed* was fixated more than once and would be expressed as a percentage value. Fixation 8 was a regression of 24 character spaces and landed back on the word *journey* meaning that when calculating the probability to regress into the word *journey* this observation would count towards the observations on which a regression from any rightward word was made into the word *journey*. Following that regression, the eyes moved forward again ending with fixation 10 on the last word of the sentence – the noun *scenery*. As shown in Figure 1.1, longer and more difficult-to-process words such as *impressed* received more and longer fixations than shorter and easier-to-process words such as *with*. In the case of the second word in the sentence, it did not even receive a fixation (i.e., it was skipped). It should be clear that the eye movement pattern provided in the figure demonstrates how fixations and saccades in reading reflect moment-to-moment cognitive processing.

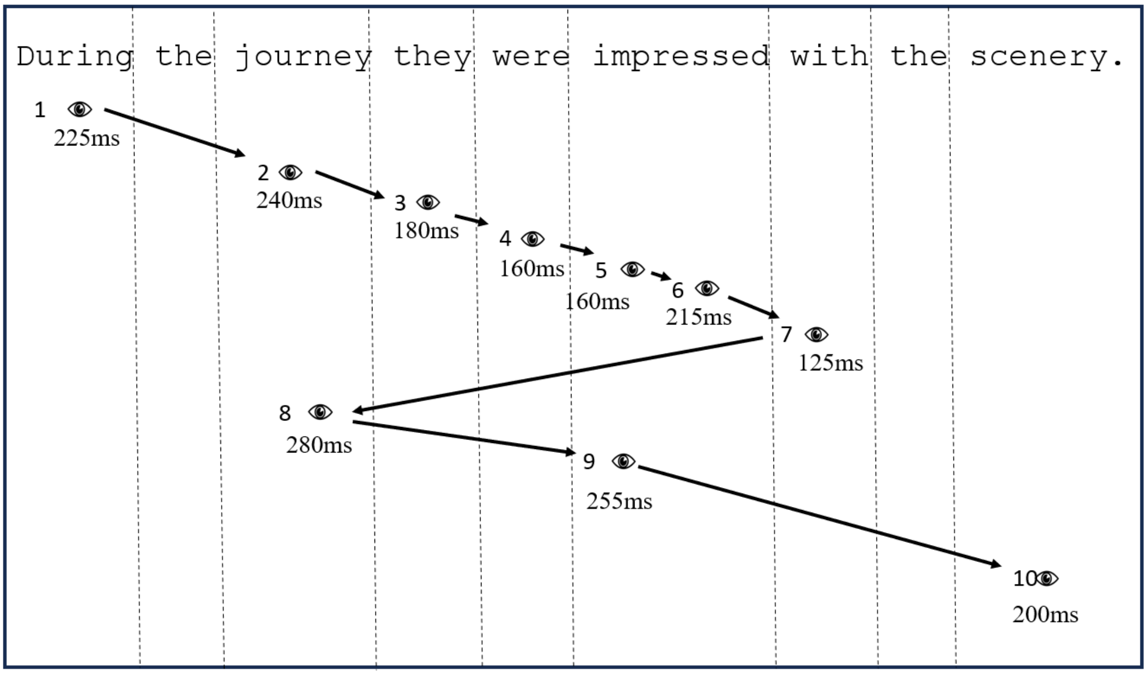


Figure 1.1. An Example set of eye movements from a hypothetical participant reading a sentence.

Each of the fixations made by the reader is represented by an “👁” symbol. The duration of each fixation in milliseconds appears underneath the fixation, and each successive fixation appears at a point lower on the page than its predecessor. Additionally, each fixation was made on a word in the sentence as denoted by the dashed vertical lines segmenting the sentence into regions such that the right boundary of each region is the right boundary of each word.

1.4. Factors that influence oculomotor behaviour during reading

In line with physiological evidence, multiple eye movement studies (e.g., Inhoff et al., 2003; Morris et al., 1990; O'Regan, 1979, 1980; Rayner, 1979; Rayner & McConkie, 1976) have shown that different variables often affect the decision of *when* to move the eyes compared with *where* to move the eyes. Multiple experimental studies have shown that visual factors (e.g., Chiu & Drieghe, 2023), as well as linguistic factors at the level of individual words (e.g., Pollatsek et al., 1999; Rayner et al., 1998) and the level of sentences and paragraphs (e.g., Hyona, 1993; Ehrlich & Rayner, 1981; Raney & Rayner, 1995; Rayner, Raney, et al., 1995; Rayner, Chace et al., 2006), affect when the eyes move. With respect to where the eyes move during reading, evidence suggests that the size (amplitude) of a saccade is primarily influenced both by the length of the fixated word, and the length of the upcoming word in the parafovea (e.g., Cutter et al., 2017, 2018; Inhoff et al., 2003; Juhasz et al., 2005, 2008), as well as the presence or absence of spaces between words (e.g., McConkie & Rayner, 1975; O'Regan, 1979,1980; Juhasz et al., 2005). Several studies have also shown that readers make shorter saccades when the upcoming word is less frequent (e.g., Li et al., 2014; Liu et al., 2016; Liversedge et al., 2014). Therefore, saccadic targeting in the context of forward eye movements may (to a lesser extent) also be influenced by the visual familiarity of upcoming words.

1.4.1. Factors that influence when the eyes move

When the eyes move during reading, is driven by three main lexical factors. These are the length of the fixated word, the lexical frequency of that word, and the predictability of the word given prior context (Clifton et al., 2016). For instance, reading times increase proportionally with word length (e.g., Just & Carpenter, 1980; Rayner et al., 1996). However, it is important to note that very short words (e.g., two to four letters) are only fixated directly 25% of the time while eleven-letter and longer words are very likely to receive more than one fixation (see Rayner, 2009). Consequently, the relationship between reading times and word length can be partially attributed to the number of fixations, not just their durations.

In contrast to word length, studies have shown that as the frequency of a word increases, reading times on that word decrease (e.g., Henderson & Ferreira, 1990, 1993; Hyönä & Olson, 1995; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kennison & Clifton, 1995; Raney & Rayner, 1995; Rayner, 1977; Rayner & Duffy, 1986; Rayner & Fischer, 1996; Rayner & Raney, 1996; White, 2008). Furthermore, when preceding context provides strong cues as to the potential word identity (i.e., the predictability of the word increases), reading times have been shown to decrease (e.g., Ehrlich & Rayner, 1981; Rayner & Well, 1996; Smith & Levy, 2013; see Staub, 2015 for a review).

Besides these three main factors, other variables at the level of the word, the sentence, and wider discourse, also influence when the eyes move. Crowding is the inhibition of letter recognition that occurs due to the proximity of adjacent letters (of the word, or other words) and their proximity to the foveated region of text. When letters appear more closely adjacent to each other (in crowded circumstances) they are more difficult to identify. Several studies to date have found that reading is disrupted by increased visual crowding (e.g., Chiu & Drieghe, 2023; Paterson & Jordan, 2010). Importantly, multiple studies (e.g., Lettvin, 1976, see Rosenholtz, 2016 for a review) have shown that mistaking letter identities and letter order in the periphery is not simply due to lower visual acuity but is also driven by crowding. In other words, while crowding may be a visual factor, it may also impact how readers process the letter make-up (orthography) of parafoveal words.

While visual factors may affect reading times there is substantial evidence from studies utilising the disappearing text paradigm (see Figure 1.2) that reading times are primarily driven by linguistic factors (e.g., Blythe et al., 2011; Liversedge et al., 2004; Rayner et al., 2003, 2011; Rayner, Liversedge, et al., 2006; see Rayner, 2009). In disappearing text studies, as each word is fixated, it disappears from the visual display after a fixed amount of time following the onset of the first fixation on that word (e.g., after 60ms). Disappearing text studies have consistently shown that even when a word is removed from the visual input, reading comprehension and overall sentence reading times are not affected (i.e., they are comparable to when text is presented normally). Furthermore, the lexical frequency of words still reliably influences reading times on that word even when the word has disappeared within the first 40-57ms after the onset of fixation (e.g., Blythe et al., 2011).

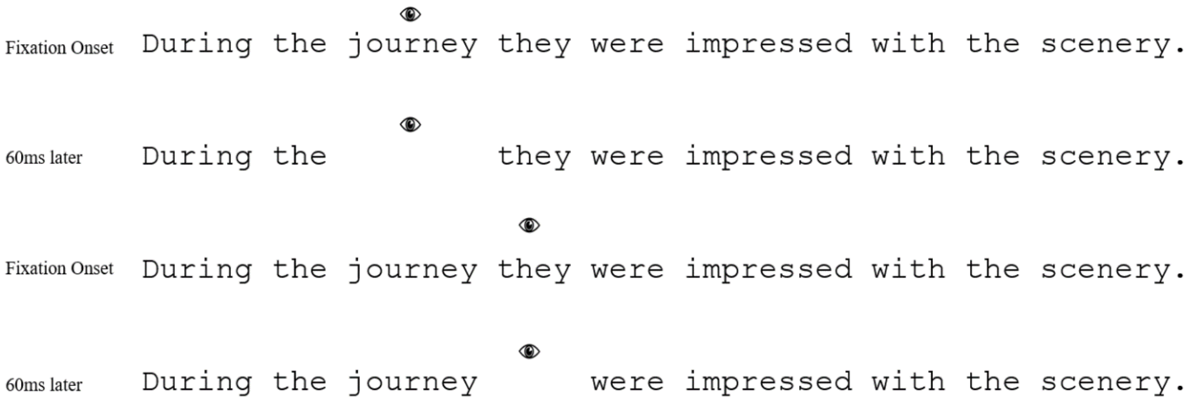


Figure 1.2. An example of the gaze-contingent disappearing text paradigm

Note: the "👁" symbol shows the position of the eye fixation. Additionally, the word underneath the "👁" symbol is the currently fixated word. As can be seen, it disappears 60ms after it is fixated.

With respect to linguistic factors at the level of words, orthographic neighbourhood size, the number of words that can be created by swapping one letter of a word to form another (e.g., *plane* - *plate*) has also been found to influence reading, with a proportional increase in viewing times as the orthographic neighbourhood size increases (e.g., Pollatsek et al., 1999) irrespective of the lexical frequency of the orthographic neighbours (e.g., Sears et al., 2006). In addition, multiple eye-tracking studies have shown that skilled readers utilise phonological information to aid word recognition (e.g., Rayner et al., 1998; see Leinenger, 2014; Milledge & Blythe, 2019 for reviews). For instance, Rayner et al. (1998) found that first-pass reading times on a target word embedded in a paragraph did not differ substantially between when the word was correct (given previous context) and when the word was incorrect but shared the same phonology, (e.g., *meet – meat*).

Besides orthography and phonology, age of acquisition, that is, when a word is learnt (e.g., Dirix & Duyck, 2017; Joseph et al., 2014; Juhasz & Rayner, 2003, 2006), has been shown to increase reading times as well. That is, readers spend longer viewing words that they have learnt more recently than words they have known for a longer time and this factor can affect eye movement measures as early as single fixation duration (Juhasz & Rayner, 2003) or the first fixation duration (e.g., Juhasz & Rayner, 2006) during first-pass reading.

A considerable number of experiments have shown that reading times of participants are influenced by their sentence parsing, as in *garden path* sentences such as *While Anna dressed the baby that was small and cute spit up on the bed* (e.g., Ferreira et al., 2001) or *The horse raced past the barn fell* (Bever, 1970). The processing difficulty in such sentences occurs because readers initially syntactically misanalyse the sentence. For example, in the first example above, a critical word (*baby*) is initially processed as an entity being dressed by Anna, however, upon encountering the phrasal verb *spit up* readers detect their initial misanalysis and realise that they need to reparse the sentence to form an interpretation in which Anna dressed herself, not the baby. When reading *garden path* sentences, participants have been found to take longer reading the critical disambiguating word or phrase than when the sentence is not ambiguous (e.g., Binder et al., 2001; Clifton et al., 2003; Frazier & Rayner, 1982; Rayner & Frazier, 1987; Rayner et al., 1992). Beyond garden path effects, there has been substantial empirical work aiming to examine how the thematic relationships between words modulate reading times (e.g., Clifton et al., 2003; Pickering & Traxler, 2003; Staub & Clifton, 2006; Traxler et al., 2002; Trueswell et al., 1994). Thematic relations represent the basic notions of what, or who, did what to what or whom (e.g., in the sentence *Andy ate* *breakfast, ate* is the verb expressing the action, *Andy* is the agent, the one performing the action, and *breakfast* is the patient, that is, what is being eaten). In particular, several studies have shown that readers are sensitive to the semantic plausibility of a word given previous context (Cohen & Staub, 2014; Filik, 2008; Patson &Warren, 2010; Rayner et al., 1983, 2004; Staub et al., 2007; Warren & McConnell, 2007, see Staub, 2015 for a discussion).

Rayner et al. (2004) examined reading times on a critical word (e.g., *carrots*) when it was a good thematic fit with the previous sentential context (e.g., *John used a knife to chop the large carrots for dinner*) versus when it was implausible (thematically possible but unlikely) given previous context (e.g., *John used an axe to chop the large carrots for dinner*) or when it was anomalous (thematically inappropriate) given previous context (e.g., *John used a pump to inflate the large carrots for dinner*). First-pass reading times on the critical word were inflated in the anomalous condition compared to the plausible (control) and implausible conditions. However, this was only evident in GD but not in FFD or SFD suggesting that when the critical word was anomalous, readers refixated it more often than in the other two conditions. In contrast, GPT and TVT only on the critical word were inflated when comparing the implausible to the plausible condition and this effect was smaller than that of the anomalous condition. This study provided a clear demonstration of thematic influences on reading times. Later studies have shown that the sensitivity to implausibility (arising due to manipulations of thematic relations) could be observed as early as the first fixation duration on the critical word and disruption increases proportionally with respect to the severity of the implausibility (e.g., Staub et al., 2007; Warren et al., 2008).

Experimental evidence further suggests that readers are sensitive to violations of world knowledge, also known as pragmatics (Braze et al., 2002; Ni et al., 1998). For example, in a sentence such as *Mosquitos will gently close while hungrily looking for their next meal* a pragmatic violation occurs in that *mosquitoes* cannot *close*, that is, the sentence describes an action that the agent cannot perform (Braze et al., 2002). Braze and colleagues (2002) compared reading times on sentences containing either a syntactic or a pragmatic violation versus control sentences which were both syntactically and pragmatically correct (e.g., *The wall will surely crack/cracking/bite after a few years in this harsh climate*). The authors initially split the sentences into six regions of interest such that each region consisted of between one to three words. The critical region of interest was the main verb in the sentence (e.g., *crack/cracking/bite*) and the following word (e.g., *after*). The first point at which participants encountered either a pragmatic or syntactic violation was always the main verb. First-pass reading times on the target region (e.g., *crack/cracking/bite after*) were found to be longer for both the syntactic violation condition (e.g., *cracking*) and the pragmatic violation condition (e.g., *bite*) in comparison to the control sentences. However, upon inspecting each of the two words separately in the critical region, only the syntactic violation inflated reading times on the main verb (e.g., *crack/cracking/bite*) while first-pass reading times on the following word (e.g., *after*) were inflated in the pragmatic violation condition suggesting that effects of the pragmatic violation had a delayed onset in comparison to effects of syntactic violation.

Several studies to date have also shown that eye movements are sensitive to higher order linguistic influences such as effects associated with quantifiers or focus operators (e.g., Filik et al., 2005, 2009; Liversedge et al., 2002; Ni, 1996; Paterson et al., 1999, 2007; see Filik et al., 2011 for a discussion). Focus operators are words that induce linguistic presuppositions with respect to discourse context (e.g., Filik et al., 2005) and thus modulate ease of processing. The scope of a focus operator is the set of words over which the operator exerts an influence (e.g., Filik et al., 2011; Jackendoff, 1972). For instance, in the sentence *Only MARY kissed John* the focus operator *Only* serves to stipulate that no one else other than Mary has kissed John. On the other hand, in the sentence *Even MARY kissed John* the focus operator *Even* establishes that multiple other potential kissers besides Mary have also kissed John (and that those potential kissers of John range according to a likelihood scale).

Filik et al. (2009) manipulated the type of focus operator that appeared in a sentence along with the relative likelihood of an event (e.g., *Only/Even the students taught by the best/worst teacher passed the examination in the summer*). Taking this example, the combination of *Only* and *the best teacher*, would implicitly suggest that all other students who were not taught by the best teacher failed the examination. The likelihood of such a scenario would, therefore, be quite high as it is natural to expect that students who received tuition by the best teacher are the most likely to pass the exam. Hence, sentences containing the combination of *Only* and *the best teacher* should be read normally. In comparison, if readers were presented with a combination of *Only* and *the worst teacher* it would be highly unlikely that the students taught by the worst teacher were the only ones to pass the exam. Consequently, the combination of *Only* and *the worst teacher* should make the text difficult to process. That is, the implication that no other students passed would be highly unlikely because all other students were presumably taught by teachers who were better at their job than the worst teacher. With respect to the combinations including *Even* as the focus operator, the expectations would be reversed. This is because a combination such as *Even* plus *the worst teacher,* is perfectly reasonable and likely, since it is possible that the exam was easy enough so students who were taught by any teacher, including the worst teacher, passed the exam. On the other hand, the combination of *Even* and *the best teacher* may be perceived as anomalous since the default expectation should be that the students who were taught by the best teacher did indeed pass the exam. Consequently, *Even* should not be necessary in such a scenario at all.

Note that in all instances of the sentence in which the focus operator was *Even*, the implication is that students taught by various teachers passed the exam. Conversely, in the sentences using *Only*, a single specified group of students taught by a specific teacher passed the exam, while all other students taught by all other teachers presumably failed the exam. In that sense, *Only* provided a stronger constraint than *Even* with respect to how readers should process the critical event (i.e., *passed the examination*). In line with these expectations, the authors found that first-pass and later reading times on the critical region when the focus operator was *only*, were longer for the unlikely than for the likely condition (*worst* versus *best teacher*) while when the focus operator was *Even*, reading times on the critical region remained comparable regardless of the likelihood of the event. Conversely, reading times on the post-critical region (e.g., *in the summer*) remained comparable (see also Paterson et al., 2007) when the focus operator was *Only*. When the focus operator was *Even* reading times on the post-critical region were inflated for the likely (e.g., *best teacher*) compared to the unlikely (e.g., *worst teacher*) event condition. These findings clearly show that readers rapidly process contrast information established by focus operators. Moreover, disruptions to such discourse processing can be detected as early as first-pass reading measures on a critical region that is incongruent with the expectations created by processing the focus operator and its scope. Based on this brief discussion it should be apparent that reading times as measured by eye movements provide a measure that is sensitive to aspects of higher-order linguistic processing during reading.

1.4.2. Factors that influence where the eyes move

In this section, I will focus on *where* the eyes move during reading. In the following paragraphs, I will first examine evidence regarding saccade length, or the distance in characters between where the saccade originated and where it landed in text (e.g., Rayner, 1998, 2009). Following that, I will focus on two notable phenomena, word skipping, the occasion when a word is not directly fixated during first-pass reading (e.g., Drieghe, 2011), and regressions, or eye movements opposite to the direction of reading (Inhoff, et al., 2019).

The choice to start with saccade length (amplitude) is purposeful because it offers the opportunity to outline basic information regarding *where* the eyes are targeted in the majority of saccades in reading, before turning to the somewhat more complex topics of word skipping and regressions.

The length of saccadic movements is variable. The length of a saccade is determined both by systematic influences as well as random noise (McConkie et al., 1988). Furthermore, saccade length is determined by character spaces in text rather than degrees of visual angle, because saccades cover approximately the same number of characters across text regardless of viewing distance, and therefore regardless of the number of letters subtending the same range of visual angle (Morrison & Rayner, 1981). The systematic component of saccade length variation has been explained as the tendency for saccades from closer launch sites to overshoot the target landing position. Conversely, when saccades are launched from further distances, they tend to undershoot the target landing position (e.g., Kapoula, 1985; Kapoula & Robinson, 1986; McConkie et al., 1988). McConkie et al., (1988) termed this observation the *saccadic range error*.

McConkie et al. (1988) showed, additionally, that there is a preferred saccade length (amplitude) in English. The authors provided evidence that when a saccade is launched from a distance of seven characters to the centre of the targeted word, the likelihood for that saccade to land at a position to the left or to the right of the centre of the word is balanced. Conversely, saccades launched from sites closer than seven characters were more likely to land to the right of the word centre, while saccades launched from more than seven characters distance were more likely to land to the left of the word centre. The notion that the saccade target is usually the centre of the word comes from studies showing that word identification is fastest when the centre of the word is fixated – an optimal viewing position (e.g., McConkie et al., 1989; Rayner et al., 1996; Vitu et al., 1990).

Evidence further suggests that saccade length may be affected by the word length characteristics of the experimental stimuli. For example, Cutter and colleagues (2017, 2018) presented readers with sentences containing words of uniform length (all the words in a sentence were three, four or five letters long) compared to sentences in which words varied in length (a mixture of words three, four or five letters long). Importantly, the authors found that as word length decreased, so did the preferred saccade length such that for three-letter words for instance, the saccade length was 4.47 characters (Cutter et al., 2018) or 4.52 characters (Cutter et al., 2017) and increased by approximately one character as word length increased from three to four and from four to five letters respectively. For sentences with variable word length, on the other hand, the authors found that the variation in saccade length associated with different word lengths was significantly reduced. Notably, Cutter and colleagues (2018) showed that change in saccade extent variability could be observed rapidly on a trial-by-trial basis and did not require repeated exposure to uniform length stimuli. Clearly, participants were able to very rapidly adapt their saccadic targeting based on the word length of the stimuli they were reading.

Multiple studies have consistently shown that forward saccades most often fall slightly to the left of the centre of a word – a preferred viewing location (PVL: e.g., McConkie et al., 1988; Rayner, 1979) instead of at the centre as per the optimal viewing position. This robust PVL effect has been replicated in multiple experiments in the last 45 years (e.g., Johnson & Starr, 2018; McConkie et al., 1989; Nuthmann et al., 2005; Nuthmann & Kliegl, 2009; Rayner et al., 1996, 1998; van der Linden & Vitu, 2016; Vitu et al., 1990, 2001; see Vitu, 2011 for a review) in alphabetic languages where the reading direction is from left to right.

Similarly to Rayner (1979), O'Regan (1981) found a convenient viewing position (CVP) which may vary across words but is on average, to the left of the word centre. O'Regan argued that word identification is most optimal when fixations fall onto the CVP, and the time needed to identify a word increases rapidly and proportionally with increasing distance of the point of fixation from the CVP. In line with this assumption, when readers fail to fixate the word slightly to the left of the centre, they are more likely to make an initial short fixation, followed by a corrective saccade to the same word (e.g., Rayner et al., 1996). Notably, this tendency to make an initial short fixation followed by a refixation is most pronounced when the eyes initially land on the word initial or word final character (e.g., Nuthmann et al., 2005). This observation has been termed the *inverted optimal viewing position* effect (e.g., Vitu et al., 2001) since the duration of the first fixation on a word is shorter when the eyes land on the first or last letter than when they land in the centre because readers are more likely to make a refixation in the former than the latter scenario.

Perhaps unsurprisingly, researchers have found that readers make a saccade to the PVL more often for longer than shorter words. In addition, where the saccade was launched from, the launch site, also reliably modulates the PVL such that launch sites closer to the word lead to saccades landing further into that word (e.g., McConkie et al., 1988; Nuthmann et al., 2005; Radach & McConkie, 1998; Rayner et al., 1996). Most studies on landing positions and the PVL, in particular, provide evidence that where the eyes land in a word is predominantly influenced by low level visual and linguistic characteristics of the text such as word length, launch site, as well as the spacing between words (Paterson & Jordan, 2010; Radach & McConkie, 1998; Rayner et al., 1998; Rayner, Yang, et al., 2013). For instance, several studies (e.g., Rayner et al., 1998; Rayner & Pollatsek, 1996) have shown that the eyes tend to land closer to the start of a word when the spaces between words are eliminated. Notably, some research has suggested that information pertaining to the letter (Hyona, 1995; White & Liversedge, 2004, 2006) or morphemic (e.g., Farid & Grainger, 1996; Hyona et al., 1989, 2018; Yan et al., 2014) make-up of a word may also influence the PVL under specific circumstances.

As noted earlier, it has been suggested that the PVL can be explained as the result of oculomotor error (a random component) and a systematic component based on the *saccadic range error* (e.g., McConkie et al., 1988). However, an alternative explanation has been proposed based on consistent findings that readers extract more meaningful information further to the right from the point of fixation than to the left given the left-to-right direction of reading in English (e.g., McConkie & Rayner, 1976; Nazir et al., 1991, 1992; Rayner et al., 1980). Consequently, landing to the left of the centre of a word is not necessarily detrimental to processing since more letters to the right than to the left of fixation can be effectively processed. Interestingly, studies in languages such as Hebrew (e.g., Deutsch & Rayner, 1999; Pollatsek et al., 1981) and Arabic (e.g., Farid & Grainger, 1996; Roman & Pavard, 1987) where the reading direction is from right to left show the opposite asymmetry in that readers obtain more information further to the left of the point of fixation than to the right. Hence, this asymmetry (termed the perceptual span, see Section 1.5.1), in the area around a fixation from which meaningful information may be extracted can be explained by the reading direction in a particular language and can in turn explain the PVL.

Recently, Johnson and Starr (2018) directly examined whether the PVL can be attributed to error in saccadic targeting or due to the increased processing of information further from the point of fixation in the direction of reading. The authors presented readers with sentences such that one word appeared per horizontal line from the top to the bottom of the screen. In this way, if the PVL is due to oculomotor error, readers should land at the centre of each word rather than to the left because the distance between launch site and landing position is uniform and the influence of word length is mitigated by the fact that the eyes need to move vertically instead of horizontally. If, on the other hand, the PVL results from increased processing in the direction of reading (i.e., from left to right within the word), then the PVL should still be observed. Johnson and Starr (2018) found that the PVL was preserved providing for the first time direct evidence that the PVL may not be explained via oculomotor error alone.

Besides where the eyes land in a word, it is important to note that approximately a third of all words are skipped during reading (e.g., Brysbaert & Vitu, 1998; Brysbaert et al., 2005; Drieghe et al., 2004, 2005; Rayner et al., 2011). With respect to word skipping, researchers have found that in English, shorter words are significantly more likely to be skipped such that three-letter words have a skipping probability of 60% while five-letter words have a skipping probability of 30% (e.g., Vitu et al., 1995). Notably, however, word length effects are modulated by where the eyes moved from (i.e., launch site – e.g., Kerr, 1992). For instance, Brysbaert and Mitchell (1996) found that two letter words were skipped approximately 70% of the time on average. When the authors accounted for the distance between launch site and the space immediately before the skipped word, they found that as distance increased, skipping decreased from 90% skipping at a one-character distance to 50% skipping at a distance of fifteen characters.

In addition, several studies (e.g., Balota et al., 1985; Drieghe et al., 2005; Ehrlich & Rayner, 1981; Rayner & Well, 1996; Schustack et al., 1987; Vitu, 1991) have shown that skipping is also affected by contextual constraint such that very predictable words are skipped more often than less predictable words irrespective of the length of that word. Furthermore, studies have also shown a smaller, but still significant, influence of word frequency on whether words are skipped (Kliegl et al., 2004; Rayner & Duffy, 1986; Schilling et al., 1998). In essence, words that appear more often in general are more likely not to receive a direct fixation than words which appear less often in text. Importantly, Brysbaert and colleagues (2005) showed in their meta-analysis that roughly only 5% and 8% of the variability in skipping in existing studies could be attributed to word frequency and word predictability respectively. Taken together, these findings show that the primary drivers of (forward) eye movements in reading are visual (e.g., spacing, word length).

Besides forward movements and word skipping, reading research has shown that approximately 5-20% of all eye movements are regressions backwards into text (e.g., Inhoff et al., 2019). Inhoff and colleagues distinguished between short and long regressions with the suggestion that these two types of regressions are guided by different factors and reflect distinct processes. Several studies looking at corpus data have shown that short regressions are typically shorter than a forward saccade with a length between one to three characters on average (e.g., Vitu & McConkie, 2000; Vitu et al., 1998) and such regressions are more likely to land at the centre of the targeted word than forward saccades (e.g., Radach & McConkie, 1998). Research suggests that short regressions to the centre of the fixated word or to the immediately adjacent prior word may serve to correct an oculomotor error, as is the case with the inverted optimal viewing position effect (e.g., Vitu et al., 2001) discussed previously. Importantly, regressions from a fixated word to the immediately preceding word are particularly common when the prior word was skipped. Similar to within-word regressions, regressions following a skip could be used to reorient the fovea to a more optimal position for the identification of the skipped word especially when skipping that word was a result of oculomotor error (see Inhoff et al., 2019). Hence, short regressions seem to aid in the word recognition process and may be affected by the processing difficulty associated with identifying a word under poor visual acuity circumstances or when an oculomotor error has occurred (e.g., when the eyes land at the end of the word or on the wrong word).

Long regressions, on the other hand, are thought to reflect a failure in comprehension of the text, that is, a misinterpretation of previously viewed information (e.g., Frazier & Rayner, 1982; see Rayner, 1998). For example, a word or sentence segment that has already been viewed and processed, may be targeted by a regressive saccade when the original interpretation of that word or text segment no longer fits with the most up-to-date sentence representation. For instance, Carpenter and Daneman (1981) showed that when readers encounter text such as *Cinderella was sad because she couldn't go to the dance that night. There were big tears in her brown dress,* they are likely to regress from the word *dress* to the word *tears* in order to correct their initially erroneous interpretation of *tears* as related to crying. Notably, there are several accounts that explain how long regressions may benefit comprehension (see Inhoff et al., 2019). Given space constraints, these accounts will not be considered in detail here. The key point with respect to the current discussion is that long regressions are usually influenced by higher order factors such as sentence parsing, interpretation, and discourse processing while short regressions seem to be driven by lower level visual factors such as a saccade landing at an unintended position in a word.

In sum, there are multiple factors which have been found to influence the time readers spend viewing a word, which word they fixate and where in that word their eyes land (see Rayner, 2009 for a review). However, a key point is that eye-tracking research investigating reading, to some extent at least, corroborates evidence from animal neurophysiological studies suggesting that there are likely two distinct systems, one of which is responsible for triggering a saccade and the other for selecting a saccade target. Crucially, evidence seems to suggest that the decision of *where* to move the eyes in terms of forward movements and short regressions (i.e., the majority of eye movements during reading) is driven by predominantly visual factors while the decision of when to move the eyes is driven by both visual and linguistic factors (although it must be noted that some factors such as word length seem to influence both decisions, see Rayner, 1998, 2009). However, notably there is significant overlap between the factors that influence fixation durations and saccade length (e.g., word length, lexical frequency, and predictability). Such overlap suggests that when it comes to such a cognitively complex and linguistically driven process as reading, the two systems that control when and where the eyes move are less independent from each other than might be assumed based on animal neurophysiological evidence. In the next sections I will focus on parafoveal processing or to what extent and which characteristics of the upcoming word or words in the parafovea may influence when and where the eyes move. Research into parafoveal processing has also been key to the ongoing debate on whether words are identified serially and sequentially or in parallel (see Zang, 2019 for a recent discussion).

1.5. Parafoveal processing

Multiple studies have shown that during a fixation, readers process not only the fixated word, but also, to an extent, the upcoming word, or words in the parafovea (see Rayner, 2009 for a review). Parafoveal processing is important to reading because it enables participants to process words before they are fixated, thus reducing the time required to process those words once they are fixated. Parafoveal processing also facilitates saccadic targeting in terms of which word is selected as the target of the next saccade and where in that word the saccade is programmed to land especially when it comes to forward saccades in the direction of reading (e.g., Rayner, 1998, 2009). Furthermore, it has been argued that efficient parafoveal processing is a fundamental aspect of skilled reading (e.g., Blythe, 2014; Tiffin-Richards & Schroeder, 2015) such that skilled (adult) readers can extract more information further into the parafovea than beginning (child) readers (Haikio et al., 2009; Rayner, 1986).

Research has shown that some visual and linguistic information is integrated across fixations (e.g., Cutter et al., 2015; Rayner et al., 1978), that is, aspects of information regarding the parafoveal word encoded during parafoveal processing is integrated with the information pertaining to the same word once it is fixated. This is termed a parafoveal preview effect, that is, a reduced reading time on the word due to it having been partially processed in the parafovea (see Schotter et al., 2012). Conversely, it is also possible that information pertaining to the word in the parafovea may influence the processing of the foveated word (a parafoveal-on-foveal effect, see Drieghe, 2011 for a review). Importantly, gaze-contingent display change paradigms (e.g., McConkie & Rayner, 1975; Rayner, 1975) have been the primary category of tool for investigating parafoveal processing. Gaze-contingent change paradigms manipulate what information is available parafoveally as the eyes move through text. The exact mechanism by which information in the parafovea is manipulated differs from paradigm to paradigm. In the next sections I will focus on the moving window paradigm (McConkie & Rayner, 1975) and the boundary paradigm (Rayner, 1975) in turn. Detailed descriptions of the mechanisms of each paradigm will be provided at the start of their corresponding section.

1.5.1. Moving window paradigm

In the case of the moving window paradigm (McConkie & Rayner, 1975; see Rayner, 2014 for a review) a window of normally presented text (of length that may be manipulated) is presented around the point of fixation such that as the fixation point moves, the window moves with it (see Figure 1.3). The moving window manipulation has been utilised to explore the extent of the perceptual span (i.e., the range of characters to the left and right of fixation from which information can be extracted during a fixation, see Pollatsek et al., 1981; Rayner et al., 1980; 1982).

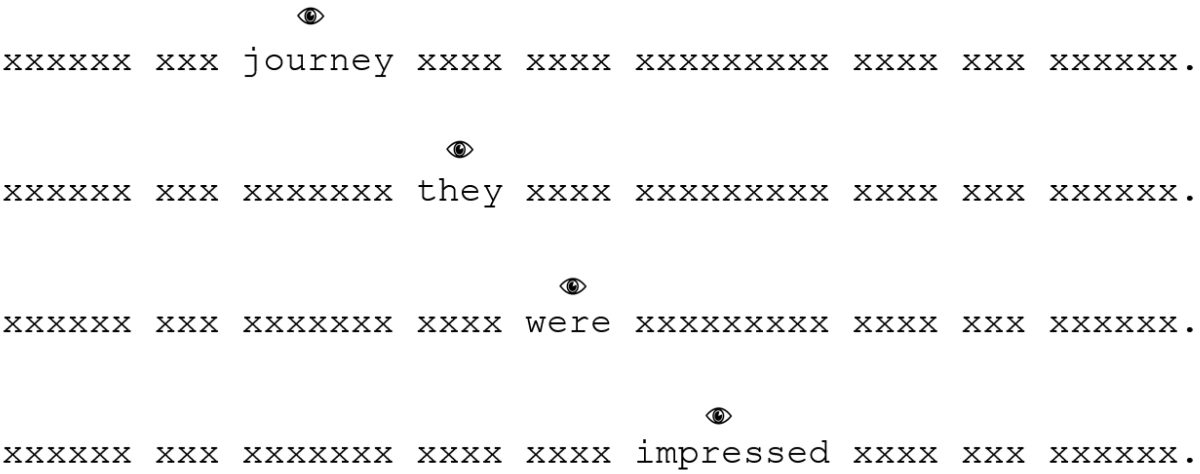


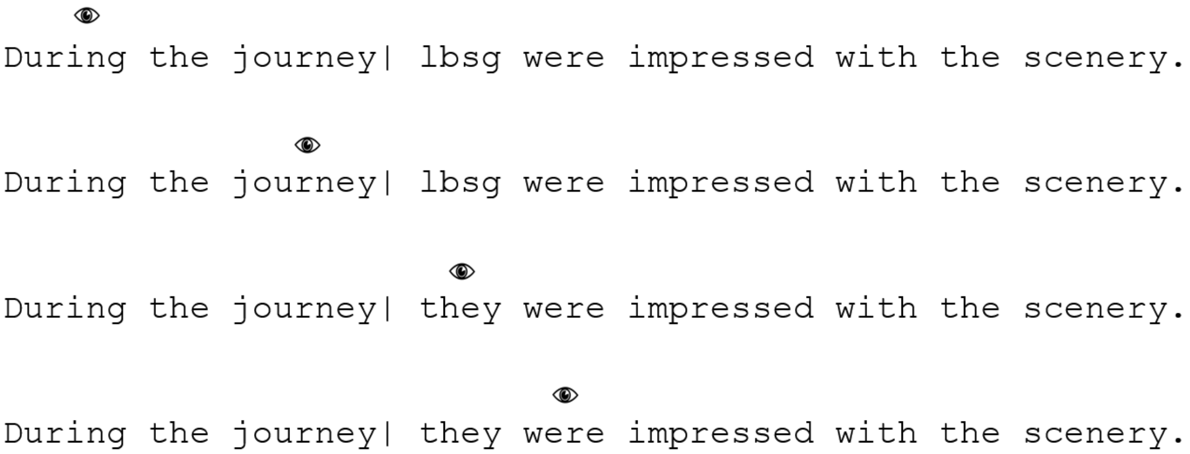
Figure 1.3. An example of the gaze-contingent moving window paradigm

Note: the “👁” symbol shows the position of the eye fixation. Additionally, the word underneath the “👁” symbol is the currently fixated word. As readers move their eyes, an invisible window moves with the point of fixation such that all surrounding words to the left and the right of the window around the point of fixation are replaced by strings of Xs.

Studies using the moving window paradigm have consistently shown that skilled English readers obtain information from approximately three characters to the left of the point of fixation out to about 15 characters to the right of the point of fixation (McConkie & Rayner, 1975, 1976; see Rayner, 2014 for a review). Importantly Rayner et al. (1982) additionally found that reading times and saccadic targeting did not differ between moving window conditions in which windows were determined in numbers of letters compared with windows formed from whole words. For example, performance was comparable when the window extended to nine letters to the right of fixation versus when the window contained the two upcoming parafoveal words. Additionally, having the whole word preserved in the parafovea did not improve reading in comparison to only having the first three letters of the word within the window. Miellet et al. (2009) further provided clear evidence that the perceptual span is influenced by attention allocation and is not simply driven by visual acuity. By enlarging the text to the left and right of the point of fixation for each fixation, the authors aimed to remove the visual acuity drop-off associated with parafoveal and peripheral vision. Despite the *magnification* of the text in the parafovea, the authors still observed a perceptual span of three characters to the left and 15 characters to the right of fixation. Moreover, multiple studies have shown that the extent of the perceptual span changes with age (e.g., Haikio et al., 2009; Rayner, 1986, 2009; Sperlich et al., 2016), the linguistic density of the orthography of the language (Inhoff & Liu, 1998; Pan & Yan, 2024; Yan et al., 2015; see Li et al., 2022 for a discussion), the differences in reading skill (e.g., Veldre & Andrews, 2014), and is asymmetric in the direction of reading of the language (e.g., Osaka & Oda, 1991; Pollatsek et al., 1981) or as specified in the task (e.g., Inhoff et al., 1989).

1.5.2. Boundary paradigm

The boundary paradigm (Rayner, 1975) allows researchers to manipulate the parafoveal preview of one or more words before they are fixated. This is achieved via an invisible boundary placed at the end of a pre-target word such that once the eyes move over the boundary on the screen (see Figure 1.4), the preview of the critical target word to the right of the boundary changes to the actual word. By manipulating the relationship between the target word in foveal vision and its preview in parafoveal vision, researchers are able to investigate what type of information (e.g., visual, orthographic, etc.) can be processed parafoveally. There has been extensive research using the boundary paradigm (Rayner, 1975) to examine what information is available parafoveally, to what extent information from the parafovea may affect reading (see Andrews & Veldre, 2019; Cutter et al., 2015; Rayner, 1998, 2009; Schotter et al., 2012; for reviews) on the fixated word (parafoveal-on-foveal effects) and to what extent parafoveal information can influence processing of the next word once it is fixated (parafoveal preview).



During the journey| lbsg were impressed with the scenery.

During the journey| lbsg were impressed with the scenery.

During the journey| they were impressed with the scenery.

During the journey| they were impressed with the scenery.

Figure 1.4. An example of the gaze-contingent boundary paradigm

Note: the “👁” symbol shows the position of the eye fixation. Additionally, the word underneath the “👁” symbol is the word that changes as the eyes move across the invisible boundary. As the readers move their eyes over the invisible boundary the word underneath the “👁” symbol reverts back from its preview form to its original form.

Multiple boundary paradigm studies have shown that the letter make-up of a word (orthography) is accessible parafoveally. In particular, there is evidence that reading times on the post-boundary word are increased when the orthography of the parafoveal preview mismatches the orthography of the word once it is fixated (e.g., Johnson et al., 2007; Pagan et al., 2021; Rayner, White et al., 2006). This cost is graded such that word-internal adjacent letter transpositions (e.g., *aksed*) in preview are the least costly, while transpositions involving word final letters (e.g., *askde* Rayner, White et al., 2006) result in longer reading times. Furthermore, when the word-initial letter is transposed in preview, this results in more disruption to reading compared to other types of parafoveal letter transpositions (Pagan et al., 2021; Rayner, White, et al., 2006). Beyond letter transpositions, several experiments have shown that when letters in a word are substituted in preview, this leads to longer reading times on that word (e.g., Johnson et al., 2007; Milledge et al., 2021; Milledge, Liversedge et al., 2022). These (and other) studies show clear orthographic parafoveal preview effects on reading times (see Rayner, 2009; Schotter et al., 2012 for reviews).

Multiple studies utilising the boundary paradigm (Rayner, 1975) to examine parafoveal processing have also shown consistently that the orthography of the upcoming parafoveal word may influence processing on the fixated word (an orthographic parafoveal-on-foveal effect: e.g., Angele et al., 2008; Inhoff et al., 2000; Starr & Inhoff, 2004). Such effects are typically observed in longer reading times on the pre-boundary word when the orthography of the post-boundary word is manipulated in preview. Drieghe et al. (2008) have suggested that parafoveal-on-foveal effects in general may be driven by mislocated fixations. That is to say, if a saccade was aimed at a word but due to an undershoot error, it landed at the end of the previous word, attention may still be allocated to the intended word rather than the fixated word. Since attention is allocated to one word but the eyes are fixating a different word, the fixation durations on the fixated word would not be indicative of processing on that word but instead they would indicate processing of the attended word. Consequently, inflated fixation durations on the fixated word would not indicate that the characteristics of the upcoming word are influencing processing of the fixated word but instead that the upcoming word itself is the one being processed.

There is also considerable evidence in favour of phonological parafoveal preview effects in reading (Ashby & Rayner, 2004; Ashby et al., 2006; Henderson et al., 1995; Jouravlev & Jared, 2018; Milledge, Zang et al., 2022; Pollatsek et al., 1992; Rayner, Sereno, et al., 1995; Tiffin-Richards & Schroeder, 2015; Vasilev et al., 2019). For instance, Ashby and Rayner (2004: Experiment 2) used the boundary paradigm to investigate whether readers are sensitive to the phonological information associated with the first syllable of the upcoming word in the parafovea. The authors manipulated the post-boundary word such that it could either start with a syllable comprising a consonant plus vowel (e.g., *de – device*) or a consonant plus vowel plus consonant (e.g., *bal – balcony*). The parafoveal preview of the post-boundary word was also manipulated such that readers could receive a preview starting with a consonant plus vowel (e.g., *de_πxw*) or a consonant plus vowel plus consonant (e.g., *bal_πxwx*) syllable while the first letter following the first syllable was masked by an underscore (placeholder) followed by a *π* and then a combination of *w*s and *x*s such that only the first syllable was available parafoveally. Hence there were two conditions in which the post-boundary word and its preview could have the same first syllable structure (either consonant plus vowel, or consonant plus vowel plus consonant) and two conditions in which there was a mismatch between the post-boundary word and its preview (preview with a consonant plus vowel first syllable and post-boundary word with a consonant plus vowel plus consonant first syllable, or preview with a consonant plus vowel plus consonant first syllable and post-boundary word with a consonant plus vowel first syllable). Analyses on reading times indicated that readers spent less time viewing the post-boundary word when its preview had the same first syllable structure versus when there was a mismatch regardless of the syllable structure, providing evidence that phonological information can be extracted from parafoveal vision.

As discussed, the existence of orthographic and phonological parafoveal preview effects as well as orthographic parafoveal-on-foveal effects (see Rayner, 2009; Schotter et al., 2012 for reviews) has been well-established. However, whether any higher order information such as the meaning or syntactic category of the parafoveal word can be processed before it has been fixated, has been at the heart of the debate on whether words are lexically processed (hence identified) in a serial or in a parallel fashion (e.g., see Reichle, Liversedge, et al., 2009; Snell & Grainger, 2019a for discussions). This debate has focused largely on three key topics. First, whether the meaning and syntactic category of a parafoveal word can influence processing on the fixated word (a semantic or syntactic parafoveal-on-foveal effect); secondly, whether the meaning and/ or syntactic category of the upcoming word may be parafoveally pre-processed thus facilitating processing on that word once it is fixated (a semantic or syntactic parafoveal preview effect), and thirdly, whether information can be extracted from two words away from the point of fixation – an n+2 parafoveal preview effect (see Andrews & Veldre, 2019; Cutter et al., 2015; Radach & Kennedy, 2013; Rayner, 2009; Schotter et al., 2012; Vasilev & Angele, 2017 for reviews). Overall, the consensus has been that if words are lexically identified in parallel during natural reading, all three of these effects should be observable and replicable. If, on the other hand, words are lexically identified serially and sequentially, then none of these effects should be evident under most natural reading scenarios (e.g., Reichle, Liversedge, et al., 2009).

The first of the three points I will focus on is the issue of n+2 preview and parafoveal-on-foveal effects. There has been limited empirical evidence in favour of n+2 preview effects (see Vasilev & Angele, 2017 for a recent meta-analysis). The meta-analysis by Vasilev and Angele (2017) only considered nine experimental studies featuring eleven experiments that investigated parafoveal-on-foveal and parafoveal preview effects associated with word *n+2* via the boundary paradigm (Rayner, 1975).

The first study to directly explore n+2 effects was conducted by Rayner et al. (2007). In their two experiments, the authors manipulated the parafoveal preview of a target word *n* such that it could either be identical to the target word in foveal vision (e.g., *carrots – carrots*), a different word (e.g., *allergy – carrots*), or a string of letters that did not comprise a word (e.g., *xonnulc – carrots*). Further, the authors manipulated the location of the boundary such that it could be placed two words (*n-2*) or one word (*n-1*) prior to the target word. The *n-2* word was always the determiner *the* across both experiments. The *n-1* word was five or six letters long and the target word *n* was between five and twelve letters long in Experiment 1. In Experiment 2, both word *n-1* and word *n* were three or four-letters long each. This change from Experiment 1 to Experiment 2 was made in order to examine whether n+2 effects, if such exist, may be easier to observe when the target word is closer to the point of fixation and falls fully into the perceptual span during the fixation on the word that is two words prior to the critical word. In both experiments, there were no reliable differences in reading times on the pre-target word (*n-1*). Moreover, when the boundary was placed at word *n-2,* there were no reliable differences in reading times on the target word regardless of the preview condition. Reading times on the target word were only inflated for the word substitution and letter mask previews when the boundary was placed at word *n-1*. Rayner et al. interpreted these findings as evidence that the linguistic properties of a word in the parafovea that is two words away in the direction of reading from the fixated word are not processed parafoveally which is inconsistent with parallel lexical processing assumptions. Later experiments by Angele et al. (2008), Angele and Rayner (2011), as well as Yang, Rayner et al. (2012) further found no evidence for *n+2* parafoveal preview effects. Specifically, Angele et al. (2008) found no evidence of a parafoveal preview n+2 effect regardless of whether the word *n+1* had a high or low lexical frequency. This finding is also inconsistent with parallel processing assumptions since if word *n+1* is highly frequent, it should require less lexical processing, thus facilitating the processing of word *n+2* in parafoveal vision.

Kliegl et al. (2007) manipulated the preview characteristics of a target word *n* via an invisible boundary set at the end of word *n-2* such that the preview of word *n* could either be identical to the target word or a letter string that did not comprise a word. Additionally, word *n+1* was always three letters long and could either be a highly frequent function word or a lower frequency content word. The authors found no evidence for a parafoveal preview n+2 effect since reading times on the target word did not change regardless of the preview condition. This finding is consistent with previously discussed investigations of the n+2 parafoveal preview effects (e.g., Rayner et al., 2007). Unlike Rayner et al. (2007), Kliegl et al. showed evidence that there was some parafoveal processing of the target word at both the pre-boundary word and the pre-target word. Specifically, both for the pre-boundary (*n-2*) word and the pre-target (*n-1*) word, reading times were inflated when readers received an invalid compared to a valid preview of the target word consistent with a parafoveal-on-foveal effect. The preview manipulation of the target word only affected gaze durations on the pre-boundary word and the effect was only observed when the pre-target word was a content word. Notably, both parafoveal-on-foveal effects were quite small (8-11ms average difference). Hence, Kliegl et al. argued that the apparent absence of significant n+2 effects could at least partially be attributed to low statistical power rather than a true null effect. Further studies by Yang et al. (2009), Yan et al. (2010), Risse and Kliegl (2012), and Radach et al. (2013) have shown evidence for small but reliable parafoveal-preview n+2 effects and parafoveal-on-foveal n+2 effects on the pre-boundary *n*-2 as well as the pre-target *n*-1 word.

In their meta-analysis, Vasilev and Angele showed that across all n+2 studies they considered, there was a small parafoveal preview n+2 effect (on average approximately 5ms). The effect was found to be larger in studies on non-alphabetic (i.e., Chinese) than alphabetic (e.g., German or English) languages. Furthermore, the n+2 preview effect was strongest on first fixation duration. In comparison, the authors found that the typical preview effect associated with a valid compared to an invalid preview of word *n+1* across 88 experiments was approximately 45ms for gaze duration. Overall, Vasilev and Angele argued that there is a small parafoveal preview n+2 effect ultimately suggesting that there is some limited parafoveal processing of word *n+2*. However, based on the existing evidence, it is not fully clear whether the parafoveal n+2 preview effects are driven by linguistic or visual factors. It may be possible that n+2 effects are purely visual and caused by the change from preview to target regardless of how the preview and target word are related. In addition, Schotter et al. (2014) showed that even if lexical processing is serial (as per E-Z Reader: Reichle et al., 1998; Reichle, 2011; see Section 1.6.1), it is possible to obtain a parafoveal preview n+2 effect if attention is shifted rapidly enough from word *n-2* to word to word *n-1* and subsequently to word *n* at the point of fixating word *n*. This could only happen if both word *n-2* and word *n-1* are identified as word *n-2* is under fixation, meaning that parafoveal preview n+2 effects should be rare and in principle should only be orthographic (and, potentially, phonological). Consequently, such effects can be reconciled with a serial processing account. If on the other hand, the n+2 effects are driven by linguistic factors and are observed consistently, then they would fit into parallel lexical processing assumptions. Therefore, empirical investigations of n+2 effects have not provided conclusive evidence in favour of either the serial or parallel lexical processing perspective.

Next, let us turn to the issue of semantic and syntactic parafoveal preview effects. At the time of publication of the review on parafoveal processing by Schotter et al. (2012), there were no experimental studies in English that had found semantic parafoveal preview effects. Yet such effects were obtained in other languages (see Radach & Kennedy, 2013 for discussion) such as German (Hohenstein et al., 2010; Hohenstein & Kliegl, 2014) and Chinese (Cui et al., 2013; Li et al., 2023; Tsai et al., 2012; Yan et al., 2009, 2012; Yang, Wang, et al., 2012). Additionally, Schotter (2013) did find a semantic parafoveal preview effect in English when comparing reading times on a target word that had a synonym preview (e.g., *rollers – curlers*) versus an unrelated word preview (e.g., *suffice – curlers*) and an identical preview (e.g., *curlers – curlers*). Schotter showed that readers spent less time viewing the target word when its preview was a synonym than when it was an unrelated word, and thus, argued that semantic parafoveal preview effects in English do occur but significant overlap in the meaning of preview and target is necessary to mitigate the influence of other factors such as the orthographic (dis)similarity between preview and target.

Subsequently, Veldre and Andrews (2016a) argued that besides the semantic relatedness difference in the previews used by Schotter (2013), they also differed in the degree to which they produced locally plausible sentences. To be precise, unrelated previews created anomalous sentence continuations while synonym previews created plausible sentence continuations. Hence, Veldre and Andrews (2016a) proposed that readers were parafoveally sensitive to the plausibility rather than the semantic relatedness of the preview per se. To test whether plausibility preview effects may be driven by semantic or orthographic parafoveal processing Veldre and Andrews (2017) manipulated the plausibility of a target word in the parafovea independently from the orthographic overlap between the target word and its parafoveal preview. This resulted in four types of previews whereby the preview could be implausible and an orthographic neighbour to the target (e.g., *rate* - *rats*) be implausible and have no orthographic overlap with the target (e.g., *sigh* - *rats*), be plausible and an orthographic neighbour to the target (e.g., *rags* - *rats*) or be plausible and have no orthographic overlap with the target (e.g., *junk* - *rats*). The authors found that first-pass reading times on the target word were shorter when its parafoveal preview was an orthographic neighbour compared to when it was an orthographically unrelated word replicating previous findings that orthographic information is processed parafoveally. Furthermore, first-pass reading times on the target word were lower when it was preceded by a plausible than an implausible preview clearly showing that plausibility is also at least partially processed parafoveally. Moreover, there were no significant interactions on any measures between the orthographic relatedness and plausibility factors. This clearly suggests that the observed plausibility effects were not orthographically driven and lends support to the notion that some semantic information may be processed parafoveally.

The third point of contention in the serial versus parallel processing debate has been the existence of semantic and syntactic parafoveal-on-foveal effects. As mentioned previously, these effects have been difficult to obtain in experimental settings (Brothers et al., 2017; Cutter et al., 2019, 2020; Schotter et al., 2012). Even so, it is the case that higher order parafoveal-on-foveal effects have been found in corpus studies and analyses of large scale datasets (e.g., Kennedy & Pynte, 2005; Kliegl et al., 2006; Pynte & Kennedy, 2006; see Kliegl, 2007). For example, Kennedy and Pynte (2005) used texts from newspaper editorials while Kliegl and colleagues (2006) used samples from over 200 participants ranging from 16 to 84 years old. In other words, such studies have utilised materials that could not be tightly controlled for visual variables such as word length and linguistic variables such as lexical frequency or predictability which are known to account for a significant portion of variance in local eye movement measures (Clifton et al., 2016; Rayner, 2009). Furthermore, multiple studies have shown that eye movement patterns change significantly with age (e.g., He et al., 2021; Kemper & Liu, 2007; Paterson et al., 2020; Rayner et al., 2009, 2010; Rayner, Reichle et al., 2006; Warrington et al., 2018, 2019) meaning that some of the variability in data from larger datasets could be attributed to age-related differences in oculomotor control between participants rather than as the presence of higher-order parafoveal-on-foveal effects.

These examples show how evidence from corpus and large-scale studies may be difficult to reconcile with findings from experimental studies where both participant and stimuli characteristics are tightly monitored and controlled to ensure that any observed variability in the data is not driven by confounding influences (e.g., see Rayner, 2009; Schotter et al., 2012). Furthermore, as outlined previously, there has been strong empirical evidence that parafoveal-on-foveal effects, in general, may be explained as the result of mislocated fixations (e.g., Drieghe, 2011) resulting from saccadic error (e.g., McConkie et al., 1988) and cannot be taken as conclusive evidence in favour of either parallel or serial processing accounts.

Overall, existing evidence on higher-order parafoveal-on-foveal effects is mixed and not sufficient to determine whether words are lexically identified in a serial or a parallel fashion. Therefore, in general, experimental findings on n+2, higher order parafoveal preview, and parafoveal-on-foveal effects seem to support the notion that words are lexically processed serially and sequentially rather than in parallel.

Importantly, the eye movement reading research outlined so far has been developed in parallel with multiple formal computational models of oculomotor control during reading (see Reichle, 2021 for a recent discussion). In the following section I will outline three influential oculomotor control computational models of reading that have been developed over the past 25 years and in the cases of the first two models, I will also outline their successor models.

1.6. Eye movement control models of reading

Computational models are formal theoretical frameworks that aim to describe, explain, and predict complex processes (such as reading) in terms of a set of critical (*core*) assumptions which are operationalised as mathematical equations (see Rayner & Reichle, 2010; Reichle 2021 for discussions). Given the complexity of the reading process, computational models of eye movement control have proven important in explaining existing evidence and driving the development of empirically testable hypotheses to improve and expand on existing theoretical assumptions.

Multiple frameworks have been put forward in the field over the last 25 years, however, the present thesis will focus on three main models that provide direct predictions regarding the research questions of each of the three experiments that I have conducted for this thesis, and which will be discussed in subsequent chapters. Additionally, while only three models will be discussed in detail, it is worth noting that in the case of two of the models, successor models that build on the previous work have been established as well. Therefore, both the original model and its successor model will be discussed in conjunction as they share multiple similarities in both computational architecture and theoretical assumptions.

1.6.1. E-Z Reader and Über Reader Models

First, the E-Z Reader model (Reichle et al., 1998; Reichle 2011), that is (perhaps) the most influential Serial Attention Shift (SAS) model in the field, and its successor model, the Über Reader (Reichle, 2021; Veldre et al., 2020), will be outlined and explained. Both E-Z Reader and Über Reader postulate that words are lexically processed and identified serially and sequentially, suggest that eye movements are preceded by shifts in attention (hence their classification as SAS models), and that processing occurs in distinct and sequential stages such that lexical processing needs to be completed for processing of the next word to commence.

E-Z Reader postulates that processing starts with an early preattentive visual stage which corresponds to the time it takes for the signal from the retina to reach at least the primary visual regions in the brain (e.g., the striate cortex – V1). Following the completion of the visual processing stage, there is the early lexical stage (*L1 - familiarity check*) during which the letter make-up, or orthography, of the fixated word, and to an extent of the upcoming word in the parafovea, is processed. This stage is affected by the length of the word (e.g., Kliegl et al., 2004; Paterson et al., 2015; Pollatsek et al., 2008), the frequency with which the word appears in language as measured by corpus studies (e.g., Balota et al., 2007; Francis & Kućera, 1982; van Heuven et al., 2014), and the predictability of the word given previous sentential context (e.g., Rayner et al., 2005; see Staub, 2015 for a review). The completion of the *L1* stage kickstarts the programming of a saccade to the next portion of text (which will be discussed shortly) as well as the late lexical stage (*L2 – full lexical access*). *L2* corresponds to the full lexical identification of the currently attended word. During this stage, the meaning of the word is activated, and it becomes available for integration into context. The time it takes to finalise the *L2* processing stage is directly proportional to the time it takes to complete processing at the *L1* stage and can never be zero. Consequently, the same variables that exert an influence on *L1*, also influence *L2*. Completing *L2* processing has two main consequences. These are that integration of the identified word begins as soon as it is identified, and attention is shifted to the next word to be processed. One important consequence of this two-stage account of lexical processing is that the more difficult a directly fixated (foveated) word is to process, the less the following word in the parafovea is processed before the eyes move to it, a foveal load effect (e.g., Henderson & Ferreira, 1990; although see also for example Veldre & Andrews, 2018b; and Zhang et al., 2019 for relevant work in Chinese). Under this account, words are not processed in a single visual episode (i.e., during a single fixation) but at the very least in two distinct episodes – prior to being fixated (in parafoveal vision) and during fixation (foveal vision). Hence, if a word is skipped during first-pass reading, and this was not the result of an error in where a saccade landed, the model assumes that the skipped word was fully identified solely on the basis of the *L1* processing stage in parafoveal vision. The integration stage (*I*) is the final linguistic information processing stage according to the most recent version of E-Z Reader (Reichle, Warren, et al., 2009). This integration stage proceeds in the background and is hypothesised to affect eye movements only when there is a failure to integrate a word (e.g., when the next word in text is lexically identified before the integration of the current word is complete). This failure to integrate a word into context results in the cancellation of any attention shifts to forward portions of text. Consequently, a new regressive saccade programme is initiated to a previously attended word to aid in the reintegration of the problematic word.

With respect to the programming of saccades to the next portion of text readers aim to process, the model suggests that there are two stages of processing, a first stage which begins as soon as *L1* is completed, and a second stage that follows from the first. Notably, the first stage of saccade programming allows for a saccade to be cancelled and a new saccade to be programmed to a different location (e.g., due to difficulties in processing) while the second stage is inevitable in that a saccade program that reaches the second stage cannot be cancelled and will occur even if there are any further disruptions to processing. Notably, the model postulates that under normal circumstances the saccade target is the closest unidentified word which is usually the word to the right of the fixated word in the parafovea (when the direction of reading is from left to right).

Über Reader (Reichle, 2021; Veldre et al., 2020) expands on the computational architecture of E-Z Reader in three major ways. Firstly, the model provides a more detailed account of word identification based on two source models – the overlap model (Gomez et al., 2008) and the Multiple Trace Memory model (Ans et al., 1998). According to Über Reader (Reichle, 2021), visual information (e.g., letter features and word boundaries) are used to create a memory probe for an attended word using the orthographic features of the word. These orthographic features are subsequently used to retrieve from memory the phonological, semantic, and syntactic information pertaining to that word based on an *echo* according to which, activation in the word identification system needs to reach a minimal threshold of similarity between the probe and a lexical item from memory.

The model utilises the letter position uncertainty principles of the overlap model (Gomez et al., 2008) to account for the finding that a letter at a given position may also activate words in which the letter appears in other positions with decreasing activation based on the distance between where the letter appears in print versus where it appears in the activated word. This account of word identification occurs in two stages – the creation and activation of the probe (corresponding to *L1* in E-Z Reader) and the *echo* resulting from the probe becoming stable (corresponding to *L2* in E-Z Reader).

Secondly, Über Reader expands on the sentence parsing mechanisms of E-Z Reader (Reichle et al., 2009) by substituting the *I* stage with a set of rules taken from three key sentence parsing models – the Activation-Based model (Lewis & Vasishth, 2005; Vasishth et al., 2008, see Jager et al., 2017 for a review), the Cue-Based Parser (Van Dyke & Lewis, 2003) and the Garden-Path model (Frazier & Rayner, 1982; 1990, see Ferreira & Qiu, 2021 for a review). Broadly, Über Reader postulates that sentence parsing is based on a set of rules which function like *if… then* conditional statements such as *If the current word is a determiner, and the phrase unit is a noun phrase, then attach the determiner to the noun phrase* (as per the Activation-Based Model). The model hypothesises that phrases are generated on the basis of retrieval from memory, meaning that to what phrase a word belongs can be affected by memories that do not directly pertain to the currently processed word, phrase, or sentence (as per the Cue-Based Parser: Van Dyke & Lewis, 2003). However, the model implements a preference for the simplest possible phrase structure such that each identified word is (initially) attached to the closest possible phrase (as per the Garden-Path model). Notably, however, Über Reader does not currently contain any assumptions directly relevant to scenarios where a sentence is mis-parsed meaning that the model cannot be used to derive predictions as to eye movement behaviour when there is difficulty with sentence parsing (e.g., when there is a syntactic violation in the sentence).

Finally, the model contains a discourse processing module that aims to explain how the meaning (semantics) of individual words alongside general knowledge of the world are utilised to form a representation of semantic context primarily derived from for larger text units such as passages (see Reichle, 2021 for a detailed discussion). The remaining assumptions of the model largely overlap with E-Z Reader especially in relation to word-by-word oculomotor control and attentional shifts during reading. Therefore, in further chapters, serial processing account predictions regarding the effects of word transpositions on eye movements (the focus of this thesis) will be derived from E-Z Reader (Reichle, Warren, et al., 2009; Reichle, 2011) since these predictions will not change for Über Reader (Reichle, 2021; Veldre et al., 2020).

Based on the discussion above, with respect to the serial versus parallel debate in eye movement control in reading, both E-Z Reader and Über Reader suggest that words are lexically identified serially and sequentially. Furthermore, integration of semantic and syntactic information across fixations should not occur under most natural reading circumstances resulting in no statistically significant higher-order (semantic or syntactic) parafoveal-on-foveal effects. In addition, since attention is allocated to one word at a time and may only shift to the next word that readers aim to process, effects pertaining to word *n+2* in the parafovea should only be possible for the visual characteristics of word *n+2* since all words in the visual field are processed to some degree during the preattentive visual stage of processing or the orthographic characteristics of word *n+2* in the event that attention can be shifter from word *n* to word *n+1* and subsequently to word *n+2* rapidly enough at the point of fixating word *n* (see Schotter et al., 2014).

1.6.2. SWIFT and SEAM Models

Following from the SAS models, the SWIFT (Engbert et al., 2002, 2005) and its successor SEAM (Rabe et al., 2024) models will be discussed as they provide the classic Parallel Gradient (PG) processing perspective, where multiple words may be lexically processed (and therefore identified) simultaneously during reading. Importantly, SWIFT (Engbert et al., 2002, 2005) focuses primarily on explaining the timing and targeting of saccades as only one of the model’s seven core assumptions directly pertains to reading times. Namely, the model postulates that lexical processing is spatially distributed such that multiple words may be identified in parallel. Consequently, the semantics and potentially syntax of up to (approximately) two words to the right of the currently fixated word may affect the reading times on that word (meaning that semantic or syntactic parafoveal-on-foveal effects might occur in reading). One issue that arises from the parallel lexical processing assumption is related to how do identified words get sorted into their correct sentential positions. In the framework of SWIFT, it is theoretically possible to identify words out of the order in which they are presented. However, the model does not provide a clear explanation as to how words that are identified out of order may be integrated into context and hence which word would be assigned to which position in the sentence. Beyond this, in a more recent implementation of SWIFT (Schad & Engbert, 2012), it was postulated that the spatial distribution of attention is adaptable. Hence, if the fixated word is more difficult to process (i.e., foveal load is increased), then attention is constrained to a smaller window, essentially resulting in serial lexical processing. When the fixated word is easier to process on the other hand, attention is distributed over a larger area, thus resulting in parallel lexical processing.

The model postulates, similarly to E-Z Reader, that lexical processing occurs in two stages. However, according to SWIFT there is an initial preprocessing phase during which activations build up. This stage is directly influenced by both word frequency and word predictability such that word frequency is inversely correlated with the amount of activation required for a word to be identified, while word predictability affects the speed at which activation builds up. Once a word is identified, the activation associated with it decreases in a second, later, lexical completion stage which corresponds to post-lexical processing or integrating the word within context.

With respect to saccadic targeting, SWIFT also adopts a two-stage approach. The first stage is labile and saccade programs at this stage can be cancelled, while at the second non-labile stage, any saccade program cannot be cancelled, and the saccade is executed. Crucially, saccade generation according to SWIFT is based on two distinct and parallel systems. The first system handles when a saccade is launched, with the model using a random timer as the fundamental basis for saccade timing. While the basis of the timing of the saccade is random, it is influenced by a preferred saccade rate which is inhibited by the difficulty of processing the fixated word (i.e., if the foveated word is difficult to process, the generation of the next saccade is delayed). How a word is selected as a target is also somewhat complex with a basic preference for the next closest unidentified word. Crucially, however, as the activation of a word becomes stronger its likelihood of being the next saccade target increases and as its activation decreases after it is identified, its likelihood of being fixated also decreases. In other words, SWIFT postulates that saccade targeting is not static but rather the selection of saccade targets evolves as words are being lexically processed.

SEAM (Rabe et al., 2024) expands on the SWIFT framework in a similar way to how Über -Reader expands upon E-Z Reader. Namely, SEAM integrates the Activation-Based model (Lewis & Vasishth, 2005) of sentence parsing with SWIFT (Engbert et al., 2005). In essence, SEAM postulates that a post-lexical processing stage is based on memory traces such that as words are identified they are parsed into *chunks* stored in memory. Each identified word produces a memory trace that is then used to select the appropriate *chunk* to bind with that word. Importantly, the model hypothesises that sentence parsing processes have a minimal influence over oculomotor control (Jager et al., 2020). The primary way in which this post-lexical processing may affect oculomotor control during reading is in triggering regressive eye movements. Specifically, when a memory *chunk* becomes reactivated, this may trigger a saccade backwards from the currently fixated word to that *chunk*. To this end, predictions regarding saccadic eye movements forward to upcoming words in reading would primarily be derived from SWIFT rather than SEAM.

The critical predictions of SWIFT and SEAM are that semantic and syntactic factors should affect parafoveal processing resulting in higher-order parafoveal-on-foveal and parafoveal preview effects as well as reliable effects associated with the parafoveal processing of word *n+2*. Although n+2 effects should be smaller than n+1 effects given word *n+1*'s closer proximity to the fovea and the drastic reduction in lexical processing rate as the distance between the point of fixation and the word increases.

1.6.3. OB1 Reader model

Besides the classic parallel gradient frameworks of SWIFT and SEAM, the more recent OB1 Reader model (Snell, van Leipsig, et al., 2018) also postulates parallel lexical processing. The model uses the principles of the open-bigram coding scheme (Grainger et al., 2014; Grainger & van Heuven, 2004; Whitney, 2001) to explain word identification such that all letters within the same word that are within three character spaces from each other activate possible two-letter combinations (bigrams) that adhere to the order in which the letters appear in the text. The bigram nodes that receive activation are then used to activate words which contain those bigrams up to a point in time when a single word from all activated word candidates is successfully identified. Importantly, the authors state that it is possible, although not implemented in the model, for open bigrams associated with a particular word to be activated on the basis of letters belonging to other different words which might in principle leading to the erroneous activation of words that are not presented in the text. Further, open bigrams are activated for the fixated word as well as up to two words to the left and two words to the right of the fixated word. The size of this attentional window in which letters, bigrams and words receive activation is adaptable, such that when a word is successfully identified, the window increases in size (five words in total) and when there is a failure in word identification, the window shrinks in size to only incorporate the fixated word as well as the word immediately to the left and the word immediately to the right. The size of the window can, therefore, change with every processing cycle of the model which is set at 25ms. While the attentional window may be symmetrical, in terms of its extent to the left and to the right of the point of fixation, letters, bigrams and words to the left of the fixated word receive less attention than letters, bigrams and words to the right in line with perceptual span evidence

(McConkie & Rayner, 1975; see Rayner, 2014 for a review). Furthermore, the word initial and word final letters are processed faster than word-internal letters due to crowding. The rate of processing outside of this attentional window is equal to the rate of processing of the word or words to the left of the fixated word. Similarly to E-Z Reader and SWIFT, word identification is primarily influenced by word length, lexical frequency, and contextual predictability.

Notably, similar to the assumption of the overlap model (Gomez et al., 2008) that letter position coding is noisy, Snell, van Leipsig, and colleagues (2018) hypothesise that word position coding is subject to uncertainty. The way that the model operationalises word position coding is via a spatiotopic sentence-level representation in working memory. This sentence-level representation is formed on the basis of word length cues as well as syntactic constraints (proposed in principle by the authors but not implemented in the model). In other words, based on word length expectations, a seven-letter word, for instance, may not be placed in the slot for a four-letter word and vice versa regardless of the order in which words are presented, viewed, or identified. This spatiotopic representation allows OB1 Reader to explain how parallel lexical processing may occur in such a manner that the meaning and potentially syntactic category of a word to the right of fixation in the parafovea does not influence processing of the currently fixated word. This is because, the information pertaining to each processed word is associated with a *blob* of an approximate length at a specific expected position in the sentence. Therefore, higher order semantic and syntactic information pertaining to each word is not integrated across words but is rather kept separate and associated with the location (*blob*) to which that word was assigned. The authors also postulate that syntactic constraints play a role in how words are assigned to their correct positions. For example, based on the identification of a noun at position 1 in the sentence, there would be a stronger expectation for a verb than an adjective at position 2 in the sentence.

Importantly, the model suggests that semantic and syntactic information from words in the parafovea may still exert an influence on the currently fixated word (a semantic or syntactic parafoveal-on-foveal effect), however, this influence is only evident during isolated word recognition via paradigms such as the flanker task (e.g., Snell et al., 2017: Experiment 2) or neurally during natural reading via brain activity measures such as electroencephalography that allows for non-invasive recordings of electrophysiological cortical activity via scalp electrodes (e.g., Snell, Yeaton, et al., 2023). Besides these assumptions regarding lexical processing and integration of words into context, OB1 Reader also aims to predict saccadic targeting. The model adopts an approach similar to the one used by SWIFT (Engbert et al., 2002, 2005) such that the target of the saccade is determined by the pooled activations of all constituent letters in a word such that longer words closer to the current point of fixation are targeted in the majority of cases. The model further postulates that whether to initiate a saccade program or not is derived from a random normal distribution. During each processing cycle, a value from that normal distribution is sampled at random and depending on where the value falls within the distribution, a saccade programme is either initiated or not initiated. Successful word identification during a processing cycle leads to more values from that normal distribution being acceptable as triggers for a saccade programme. Conversely, failure to identify a word results in fewer values from the normal distribution being acceptable as triggers to initiate a saccade programme.

1.6.4. Interim summary and discussion

Overall, the oculomotor control models discussed in this section aim to predict several aspects of eye movements during reading. Namely, the durations of fixations, the selection of saccade targets and the decision of when to launch a saccade. All the models discussed here agree that lexical processing occurs in stages and many eye movement patterns can be reliably explained primarily via a set of low level linguistic variables such as lexical frequency, word length and word predictability. Finally, the models discussed so far make limited predictions regarding higher order syntactic, pragmatic and discourse processes and how these may affect oculomotor control during reading primarily because in the case of E-Z Reader, SWIFT and OB1 Reader, higher order processes are simply not directly incorporated into the architecture of each model. With respect to Über Reader and SEAM, the implementation of higher order processing assumptions is still in development meaning that both models make only limited predictions on how higher order processing may influence oculomotor behaviour during reading.

An important distinction between these models, however, concerns how they treat processing in the parafovea. Both E-Z Reader and Über Reader function under the assumption that lexical identification is strictly serial and sequential, while SWIFT, SEAM and OB1 Reader all posit that multiple words are lexically processed in parallel, and can thus, be identified out of order and simultaneously. However, unlike the SWIFT and SEAM models, OB1 Reader does not stipulate that semantic and syntactic information pertaining to one word would affect the processing of another word with respect to local eye movement measures, hence why OB1 Reader would make predictions regarding semantic and syntactic parafoveal-on-foveal effects more akin to those generated by both E-Z Reader and Über Reader than any of the other parallel gradient models.

1.7. Sub-lexical and non-alphabetic transposition effects in reading.

As discussed previously in section 1.5.2 there has been a significant amount of research using eye-tracking and, in some instances, the boundary paradigm (Rayner, 1975) to examine, in alphabetic spaced languages, whether there is any flexibility in how readers encode letter positions during reading. As noted, results from such eye-tracking studies tend to show that letter transpositions can produce quite variable disruption to processing depending on a range of factors such as whether the two transposed letters are adjacent or not (e.g., Blythe et al., 2014; Johnson, 2007, 2009; Johnson et al., 2007; Pagan et al., 2021; Rayner, White et al., 2006; White et al., 2008). However, while processing may be disrupted, readers are still able to comprehend sentences when letters within words are transposed, suggesting that they are still able to extract information from such words and form coherent sentential representations despite disruption to processing at the point of encountering the letter transposition. While letter transpositions in alphabetic spaced scripts themselves may have received considerable scientific attention, transpositions at the level of other sublexical units such as phonemes and morphemes remain relatively under-investigated, and so have, for the most part, transpositions in nonalphabetic languages such as Chinese and alphabetic unspaced languages such as Arabic and Hebrew. In this section, before delving into the *Transposed-Word* effects which are the focus of the present thesis, I will provide a brief overview of the few existing studies to date that have examined transpositions of sub-lexical units in alphabetic and nonalphabetic scripts. This is an important discussion point for two reasons. First, while evidence is currently limited, such research may show whether transposition effects exist across different kinds of linguistic units and may not be specific to a particular language. Secondly, by limiting focus to studies examining processing of letters means that consideration of processing beyond such relatively shallow levels is neglected. Clearly readers do undertake linguistic processing at levels beyond letters during natural reading (e.g., phonemes and morphemes), and therefore, it is likely that transposition effects might occur and have implications for such processing. Furthermore, in nonalphabetic languages, the linguistic units or constituents over which transposition effects might occur, are quite different in respect of their orthographic status, and indeed, in some languages (e.g., Chinese), there can even be significant ambiguity as to the orthographic units that comprise a word (e.g., Zang, 2019).

The first point of focus in this section will be letter transpositions in alphabetic unspaced languages. Some eye-movement studies in Thai, an unspaced alphabetic language (e.g., Winskel & Perea, 2013; Winskel et al., 2012), have shown some comparability in letter transposition effects to alphabetic spaced languages. Conversely, studies on semitic languages (e.g., Hermena et al., 2021; Velan et al., 2013) have shown that letter transpositions, particularly in morphologically complex words, cause significant disruption to processing as they are encountered in text.

Semitic languages differ from most other alphabetic scripts in that they are unspaced and read from right to left. Furthermore, in both Arabic and Hebrew, vowels are usually omitted from written text under normal circumstances unlike in most alphabetic spaced languages such as English. In semitic languages, most words of semitic origin are morphologically complex as they include both a root (containing the meaning of the word) and a pattern (containing auxiliary information). The root itself is comprised of three consonants which do not need to be adjacent – are not concatenated (see Shimron, 2006). Given all these interesting characteristics, it is important to investigate whether effects of letter transpositions in semitic languages may be comparable to such effects in other alphabetic scripts.

Velan et al. (2013) first explored how native Hebrew speakers were able to read sentences in which two internal adjacent letters within a critical word could be transposed or appear in their correct order. There were three types of transpositions used in the study. First, in some cases two of the root letters were transposed creating a pseudoroot (carrying no meaning). Alternatively, transposing the two root letters in a second condition resulted in a different existing and unrelated root. Finally, letters could also be transposed within morphologically simple words which do not contain the characteristic root and pattern structure. The authors found that the first fixation duration on the target word was only inflated when the letter transposition resulted in a pseudoroot but not when it resulted in an unrelated root or when the word itself was simple. Conversely, both types of letter transpositions involving root letters, as well as transpositions in simple words, resulted in significant disruption to reading on the target word for later measures (e.g., GD and TVT). Hence, these results support the notion that letter transposition effects may not be universal, and instead may be dependent on the properties of the specific script under investigation.

In a recent study (Hermena et al., 2021) investigated how letter transpositions in preview affect reading in Arabic. They manipulated the parafoveal preview of a critical complex target word. Two of the letters in the root of the word could be transposed in preview resulting in either a new root or a pseudoroot which was meaningless. Additionally, the target word could be substituted by a synonym or an unrelated word or remain the same in preview. Finally, the preview could retain only the root or the pattern morpheme of the target word overall resulting in seven conditions. Notably, for the two transposed-letter preview conditions, the transpositions could involve adjacent letters, or letters that were up to two letters away from each other. Despite this variability, the authors found that both transposed-letter conditions resulted in significant and early disruption (i.e., FFD, SFD and GD) on the target word. Reading times for the transposed letter preview condition resulting in a new root were increased compared to the unrelated condition. Reading times for the unrelated preview and the transposed letter preview condition resulting in a pseudoroot were, conversely, comparable. Further, reading times were inflated for the unrelated and the transposed-letter pseudoroot compared to the identity preview condition. These results are important because they are a further, clear, indication that effects due to processing associated with letter transpositions differ across languages.

Beyond letter transpositions, there have been no eye-tracking studies investigating reading that have investigated potential effects of syllable and phoneme transpositions.

Additionally, there has been only a single eye-tracking study to date that has investigated the effects of morpheme transpositions on reading. Angele and Rayner (2013) manipulated the parafoveal preview of a complex bimorphemic target word via the boundary paradigm (Rayner, 1975) in two experiments. In the first experiment, the target word could remain identical (*cowboy* -> *cowboy*) or the two morphemes in the word could be transposed in preview (boycow -> cowboy). In the second experiment, the authors also included a second factor such that each morpheme could also be masked or remain intact in preview The first morpheme could be masked by a string of letters in preview (e.g., *enzboy*-> *cowboy*). The second morpheme could be masked by a string of letters in preview (e.g., *cowtxg* -> *cowboy*), or both morphemes could be masked in preview (*enztxg* -> *cowboy*). As early as FFD on the target word, reading times were influenced by all preview changes. Specifically, first-pass reading times on the target word were significantly longer when the two morphemes were transposed in preview versus when the preview was identical across both experiments. Further, if at least one morpheme was left intact and both morphemes were presented in their correct order in preview reading times were shorter than when both morphemes were masked in preview. The authors suggested that at least for relatively short complex words (six or eight letters long) in English, readers are potentially able to process both morphemes simultaneously.

Beyond work on transpositions in alphabetic languages, it is important to note that several studies to date have examined transpositions in Chinese, a logographic unspaced language (e.g., Chang et al., 2020; Gu et al., 2015: Experiment 2; Yang, 2013) using the boundary paradigm (Rayner, 1975). Specifically in these studies, the order of two characters forming a two-character word could be transposed in parafoveal preview or remain intact in preview. Furthermore, several studies, as will be discussed later, have examined word transpositions in Chinese during standard reading (e.g., Liu et al., 2020, 2021, 2022, 2024). I am mentioning these studies here, because in all of these four studies, two, single-character words were transposed to create an ungrammaticality. Note that none of these studies used eye tracking methodology, and all of them will be discussed in the subsequent section. Hence, while these studies are framed as investigations of word transpositions, they can also be viewed as investigations of character transpositions in Chinese. This again illustrates the ambiguity that exists concerning what orthographic units constitute a word in Chinese (see Zang, 2019).

In her study, Yang manipulated the parafoveal preview of a two-character target word such that the two characters could appear in their correct order, out of order, or be substituted by two characters forming an unrelated word in preview. Specifically, the transposition of the two characters could result in a synonymous word, meaning that the transposed preview did not alter the meaning of the upcoming word (Experiment 1). Alternatively, transposing the two characters could result in a different word with a different meaning (both Experiment 1 and Experiment 2). Further, in Experiment 2, the transposed preview could result in a plausible or implausible word given previous context.

In the first experiment, reading times on the target word were shortest for an identity preview condition – a standard identity preview benefit. Also, when the target word was transposed in preview and that transposition retained the meaning of the target, reading times were comparable to when the preview was identical in Experiment 1. When the transposed preview resulted in a different word, reading times were substantially longer than for an identity preview across both experiments. Moreover, reading times were substantially inflated for the substitute preview compared to the transposed preview (which, recall, resulted in a different word in both experiments). In Experiment 2, it was further visible that when the target word was transposed in preview resulting in a different word, and that word was still plausible given preceding context, reading times on the target word were comparable to when it was left intact in preview.

Before considering the implications of these experiments, I will first detail two other related studies. Yang's findings have been further confirmed and expanded upon by two subsequent studies (Chang et al., 2020; Gu et al., 2015). Specifically, Gu et al. (Experiment 2) also manipulated the preview of a two-character target word embedded in a sentence such that it could be identical, the two characters could be transposed, or the word could be substituted by an unrelated word. The target word could be a simple monomorphemic word, or a bimorphemic word. The reading times on the target word were longer for a substitute compared to a transposed preview. This cost was comparable for both simple and bimorphemic words. Further, reading times on the target word were longer for the transposed compared to the identity preview conditions regardless of the word type. Chang et al. applied the same parafoveal preview manipulation as Gu et al. (2015) and Yang (2013). They manipulated the predictability of the target word such that it could be highly predictable or highly unpredictable. The authors found the same pattern of effects as observed previously. That is, shortest reading times on the target word for identity previews, longer reading times for the transposed preview and longest reading times for the substitute preview. These effects were observed both when the target word was predictable, as well as when it was unpredictable.

Overall, the results from these three studies clearly suggest that character order may, to a degree, be processed flexibly in Chinese. Also, readers likely extract some semantic information from parafoveal vision since changing the order of the two characters that form a word in preview was only disruptive when the resulting word was implausible given previous sentential context.

Based on the studies discussed in this section it should be clear that at least in some languages there is some degree of flexibility with respect to how readers encode the order of sub-lexical units. This is evident from letter transposition studies in alphabetic spaced languages (e.g., Blythe et al., 2014; Johnson, 2007, 2009; Johnson et al., 2007; Pagan et al., 2021; Rayner, White et al., 2006; White et al., 2008) and some alphabetic unspaced languages – Thai (e.g., Winskel & Perea, 2013; Winskel et al., 2012) but not others – Arabic (Hermena et al., 2021) and Hebrew (Velan et al., 2013). Further, this flexibility may also apply to morpheme order processing at least in English (e.g., Angele & Rayner, 2013) and character order processing in Chinese (e.g., Chang et al., 2020; Gu et al., 2015; Yang, 2013). Further, it should be evident that more research is necessary to examine the effects of linguistic transpositions on eye movement patterns during reading especially for sub-lexical units other than letters and beyond alphabetic spaced languages. Hence, as of now, it is clear that there is variability with respect to the degree of flexibility that exists in processing the order of sub-lexical units across languages. It is less clear how such processing may be modulated by language specific properties. In the following section, I will focus on research investigating whether readers are able to flexibly encode the order of lexical (i.e., word), rather than sub-lexical, units.

1.8. *Transposed-Word* effects

This section will provide an overview of the existing research on word transpositions and their effects on reading. Multiple studies in the past six years have investigated the influence of word transpositions primarily with respect to a variety of decision task paradigms (Dufour et al., 2022; Hossain & White, 2023; Huang & Staub 2021a, 2022, 2023; Liu et al., 2020, 2021, 2022, 2024; Milledge et al., 2023; Mirault et al., 2018, 2020, 2021; Mirault, DeClerk, et al., 2022, 2023; Mirault, Vandendaele, et al., 2022, 2023; Pegado & Grainger 2019, 2020, 2021; Snell & Grainger, 2019a, 2019b; Snell, Mirault, et al., 2023; Spinelli et al., 2024; Tiffin-Richards, 2024; Wen et al., 2019, 2021a, 2021b, 2022, 2024 see Huang & Staub, 2021b for a review). In the following pages I will summarise the existing research and outline the main theoretical explanations of *Transposed-Word* effects in the context of both empirical findings and computational models of eye movement control during reading.

Word transposition studies have become the primary way to test the hypotheses derived from the OB1 Reader model (Snell, van Leipsig, et al., 2018). Beyond their significance for the theoretical framework of OB1 Reader, *Transposed-Word* effects represent a significant aspect of reading, namely how readers sometimes fail to detect violations which should render the text they are reading difficult or even impossible to interpret in a meaningful manner (Huang & Staub, 2021b). Instead of observing near-perfect detection rates of such violations, studies have consistently shown that readers sometimes fail to detect word transpositions and instead view sentences containing word transpositions as grammatical. Notably, this failure to detect the ungrammaticality created by transposing two words in a sentence seems to occur at a rate significantly above chance since readers make reliably more errors for sentences containing a word transposition alone than when judging the grammaticality of sentences containing a word transposition and an additional syntactic violation (Mirault et al., 2018, 2020). While most investigations to date have focused on the detection rates of word transpositions, much less is known about how such word transpositions actually affect eye movement patterns during reading. The two existing eye-tracking studies on *Transposed-Word* effects to date will be discussed in depth further into this section. As discussed previously, eye movements are excellent indicators of moment-to-moment cognitive processing during reading. Hence the three eye movement studies that comprise my Ph.D. project are aimed to expand on current understanding of how word transpositions affect online cognitive processes as readers encounter word transpositions beyond the ultimate failure or success in detecting the presence of word transpositions.

The first point of discussion will be word transpositions and the grammaticality decision task (GDT). When completing a GDT experiment, participants are required to make an overt response regarding a visually presented stimulus. As the name suggests, when participants are asked to make a grammaticality decision, they need to evaluate whether or not the string of words they are presented with forms a grammatically correct sentence. For instance, the sentence *The white was cat big* would be judged as grammatically incorrect while a sentence such as *The white cat was big* would be considered grammatically correct. To date, twenty two studies have used a GDT (Hossain & White, 2023; Huang & Staub, 2021a, 2022, 2023; Liu et al., 2020, 2021, 2022, 2024; Milledge et al., 2023; Mirault et al., 2018, 2020; Mirault, Declerck et al., 2022; Mirault, Leflaec, et al., 2022; Mirault, Vandendaele, et al., 2022, 2023; Snell & Grainger, 2019b; Snell & Melo, 2024; Spinelli et al., 2024; Tiffin-Richards, 2024; Wen et al., 2021a, b, 2024). However, the critical comparison in multiple transposed-word GDT studies has not been between correct and incorrect sentences but instead between two types of incorrect sentences. For instance, Mirault et al. (2018, 2020) compared performance on the GDT for sentences such as *The white was cat big* versus sentences such as *The white was cat slowly* where besides the transposition of the third and fourth word, there is also a syntactically illegal final word. Participants were required to make speeded grammaticality decisions. Besides the two types of ungrammatical sentences participants were also presented with an equal number of grammatically correct sentences as per standard decision task procedures. The critical target words were always the third and fourth word in the sentence and since both studies shared the same materials, those words were not controlled for length, lexical frequency, or predictability. Controlling for these three linguistic characteristics is important as has been discussed previously (e.g., Section 1.4.2) since decades of research have shown that shorter, more frequent, and more predictable words require less processing, and are more likely to be skipped than longer, less frequent and less predictable words (e.g., Clifton et al., 2016). The idea of the comparison between the transposition only condition and the condition with a transposition plus the final word being ungrammatical, stems from the notion that the first sequence can be resolved into a grammatically correct sentence if the two transposed words are moved while the second sequence cannot be resolved into a grammatically correct sentence regardless of any possible word order adjustments. Furthermore, in their 2020 study, Mirault and colleagues also recorded participants' eye movements and these eye movement results will be discussed in detail later when all the studies examining the effects of word transpositions on eye movement patterns will be considered together.

These two studies (Mirault et al., 2018, 2020), and other studies, have shown that performance for sentences containing a word transposition and a syntactically illegal final word was better (i.e., higher accuracy and lower response times - RTs) than for sentences that contain a word transposition only (e.g., Hossain & White, 2023; Mirault et al., 2018, 2020). To date, the *Transposed-Word* effect has been documented in Chinese (Liu et al., 2020, 2021, 2022, 2024), French (e.g., Mirault et al., 2018, 2020), Dutch (Snell & Melo, 2024), German (Tiffin-Richards, 2024), and English (e.g., Hossain & White, 2023; Huang & Staub, 2021a; Milledge et al., 2023; Spinelli et al., 2024). Liu et al. (2021: Experiment 1) in particular examined Transposed-Word effects in Chinese using the same type of materials as Mirault et al. (2018) and showed that both when grammaticality decisions are speeded and unspeeded, readers still sometimes fail to detect the word transposition in sentences containing a word transposition only (transposed-word sentences), significantly more often than for sentences containing both a word transposition and a syntactically illegal final word (control sentences). Notably, however, the authors showed that error rates for transposed-word sentences were on average approximately 40% for the speeded grammaticality judgment trials versus approximately 20% for the unspeeded grammaticality judgment trials. Importantly, speeded grammaticality decision times were shorter by approximately 330ms than unspeeded decision times for both the sentences containing a word transposition only, as well as the sentences containing a word transposition and a final ungrammatical word. These results suggest that while reading speed does not modulate the magnitude of the *Transposed-Word* effect, readers perform more accurately on unspeeded than speeded trials. It is notable that the authors did not find a significant interaction between the type of sentence (a word transposition only versus a word transposition and a syntactically illegal final word) and the task (unspeeded versus speeded GDTs). Instead, each factor produced distinct effects.

In their second experiment, Liu et al. (2021) manipulated the placement of the word transposition in the sentence such that it could involve the third and fourth word, thus meaning that participants had some preceding context in the form of the first and second words of the sentence. Alternatively, the transposition could involve the first two words in the sentence, meaning that participants had no sentential context prior to encountering the word transposition. The authors found that accuracy decreased, and RTs increased on the GDT when the word transposition involved the third and fourth words compared to when the word transposition involved the first and second words specifically for the sentences containing a word transposition only. In contrast, for the sentences containing both a word transposition and a final ungrammatical word, the placement of the word transposition did not change the accuracy on the GDT. However, transposing the first and second words in the sentence resulted in RTs approximately 400ms shorter than when transposing the third and fourth words. That is, the authors observed an interaction between the type of sentence (i.e., containing a word transposition only versus a word transposition and a syntactically illegal final word) and the placement of the word transposition (first and second words versus third and fourth words transposed). This interactive pattern of effects led the authors to suggest that the amount of context prior to encountering a word transposition in a sentence may modulate the *Transposed-Word* effect.

Huang and Staub (2022), on the other hand, found no compelling evidence that global (Experiment 1), or local (Experiment 2) contextual information modulates the *Transposed-Word* effect. In their first experiment, the authors created two sentence paragraphs such that in the second sentence, two critical words could be transposed or appear in order and the cloze probability of the second of the two words could be high or low. Participants were required to perform a GDT. There was a numerical decrease in accuracy when the second word in the transposed word pair was highly predictable in comparison to when it was unpredictable, however, that effect was not significant. In the second study, Huang and Staub manipulated the log-transformed frequency of co-occurrence of three-, and four-word sequences up to and including the word transposition. Further, the authors included sentences with a word repetition or a word omission, as well as grammatically correct sentences to explore the failure to detect different types of syntactic violations. The three types of syntactic violations were investigated separately. The authors found that the *Transposed-Word* effect on accuracy for the GDT did not change regardless of whether the three and four-word sequences had a high or low frequency of co-occurrence. Consequently, the type of context preceding a word transposition, does not seem to modulate the *Transposed-Word* effect.

Mirault, Leflaec, et al. (2022) conducted two online experiments. In the first experiment participants were presented with grammatically correct sentences or sentences containing a word transposition always involving two internal and adjacent words. The target word pair could be orthographically related (i.e., be orthographic neighbours with a single internal letter differing between the two words) or unrelated. In the second experiment, there were two factors again. Similar to Experiment 1, two adjacent internal words could be orthographic neighbours or could be orthographically unrelated. In addition to this manipulation, the authors included three types of sentences. First, there were grammatically correct sentences which were included to ensure participants faced a meaningful decision in relation to whether a sentence was, or was not, grammatical. Second, there were sentences containing a word transposition only, and third, there were sentences containing both a word transposition and a final ungrammatical word. In the second experiment, the comparison of interest was between sentences containing a word transposition only versus sentences containing both a word transposition and a syntactically illegal final word. The authors found that accuracy on the GDT decreased, and RTs increased for sentences containing a word transposition only versus sentences containing both a word transposition and an ungrammatical final word irrespective of whether the two target words were orthographic neighbours or if they were unrelated. Taken together, these findings clearly suggest that word transposition effects are not primarily driven by contextual (i.e., how often preceding words may appear together as measured by corpus data) or orthographic factors. Mirault and colleagues have argued that the drop in performance for sentences containing a word transposition only is due to parallel lexical processing and noisy word position coding and they argue that this fits with the OB1 Reader (Snell, van Leipsig, et al., 2018). The primary argument for this line of reasoning has been that since word positions are assigned independently from word identification in the OB1 Reader framework, sentences which may be grammatical if the position of words is altered are more likely to be misjudged as being grammatical than sentences that cannot become grammatical regardless of word order assignment. Recall that word order is computed according to a spatiotopic representation based on word length and syntactic likelihood in OB1 Reader.

Several studies have also shown a *Transposed-Word* effect with a serial presentation paradigm whereby each word is presented one at a time on the screen (e.g., Hossain & White, 2023; Liu et al., 2022; Mirault, Vandendaele, et al., 2022; Milledge et al., 2023 Experiments 2 and 3; Spinelli et al., 2024). In several of these studies, five-word long sentences were presented one word at a time such that the third and fourth words could either be transposed, or not transposed at the same central position (Hossain & White, 2023; Liu et al., 2022; Mirault, Vandendaele, et al., 2022; Milledge et al., 2023 Experiment 2; Spinelli et al., 2024: Experiment 2). Conversely, in some studies, the words were presented one at a time in their respective locations (Milledge et al., 2023 Experiment 3; Spinelli et al., 2024: Experiment 2).

These effects have been observed in Chinese (e.g., Liu et al., 2022), English (e.g., Hossain & White, 2023) and French (Mirault, Vandendaele et al., 2022).

The serial presentation of words should inhibit the spatiotopic coding mechanisms proposed by OB1 Reader for two key reasons. Firstly, when words are presented serially, a spatiotopic representation of the whole sentence should be, at best, difficult to create since there could be no length cues or syntactic expectations generated for any of the words beyond the one being viewed. Secondly, as words are presented one at a time, it is impossible to process multiple words in parallel because of the absence of parafoveal information during the viewing period of each word. For these two reasons, readers should be better at detecting word transpositions when sentences are presented word-by-word instead of naturally.

Mirault, Vandendaele and colleagues (2022) have argued that the presence of a word transposition effect in serial presentation paradigms can still be harmonised with a parallel processing account. In particular, since OB1 Reader suggests that word order is determined by both word length and syntactic expectations, a *Transposed-Word* effect when words are presented serially is possible and may be driven purely by syntactic expectations. Notably, while the model hypothesises that syntactic expectations play a role in word position coding, it does not provide a clear mechanism as to how that should occur. Furthermore, as per OB1 Reader syntactic information pertaining to each word is processed in parallel. Hence, syntactic expectations should have a diminished influence in serial presentation modes where only one word can be processed at a time and syntactic expectations for upcoming words cannot be developed to the same extent. In line with this argument, the authors found that the effect of word transpositions on the GDT in serial presentation procedures only occurs in respect of accuracy, but not RTs.

Snell and Melo (2024), further failed to find an effect of word transpositions on either accuracy or RTs when sentences were presented serially in Dutch. However, in their experiment, sentences only consisted of four instead of five or more words and while word transpositions always involved two adjacent words, the positions of the two words varied across sentences such that, for example, some transpositions involved the sentence initial word while others involved sentence internal words. Since word position coding is partially driven by syntactic expectations according to OB1 Reader, it stands to reason that the shorter a sentence is, the less syntactic expectations can be developed since fewer words need to be parsed together into a sentential representation (see also Liu et al., 2021: Experiment 2). Therefore, it may be easier to detect word transpositions in four-word sentences than in five-word sentences in general which may partially explain why the findings by Snell and Melo differ from previously reported results (e.g., Hossain & White, 2023). Additionally, Snell and Melo used a different control ungrammatical condition to previous research. All previous serial presentation mode studies used the transposed-word sentences and changed the final word for a syntactically illegal substitute (e.g., *The black was bear quietly*). Snell and Melo instead created their control ungrammatical sequences by combining the first half of one grammatical stimulus with the second half of another grammatical stimulus (e.g., *the man dog here*). Hence it is unclear whether the findings observed by Snell and Melo are inconsistent with previous research in part because of the significant methodological differences. Crucially, however, Hossain and White (2023) demonstrated an effect of word transpositions on RTs even with a serial presentation when they matched the rate of presentation to the reading speed of each participant. Therefore, *Transposed-Word* effects when words are presented one at a time remain difficult to reconcile with the OB1 Reader model. Overall, serial word presentation findings, therefore, strongly suggest that the failure to detect word transpositions is not driven by parallel processing with noisy word position coding. Given this, Liu et al. (2022) suggested an alternative explanation for the effects of word transpositions. Namely, the *noisy-channel* account (e.g., Gibson et al., 2013). According to this view, the transposed words are initially detected and cause early disruption, but the language system manages to fix the issue by assigning words to their correct positions thus mitigating any later disruption.

There is also some evidence that word transpositions lead to a drop in performance within experiments adopting two other types of tasks. Firstly, Pegado and Grainger (2019) showed that readers fail to detect that two successively presented word sequences are different in a same-different matching task, more often when the difference is created by a word transposition than when it is created by a word substitution (see also Pegado & Grainger, 2020). In other words, when participants are presented with two sentences and asked to judge if the second sentence was the same as the first, they struggle more when the difference between the first and second sequence is created by transposing two adjacent internal words than when the difference is created by substituting the same two words with unrelated words. Furthermore, this effect of transpositions was stronger when the sequences were formed from valid words compared to pseudowords. The authors took this as further evidence that word position coding is noisy as per OB1 Reader and that the flexibility in encoding word positions is not task specific.

Wen et al. (2022) further found that readers are better at recalling a word belonging to a sequence containing a word transposition only (e.g., *He the throws glass there*) versus a sequence in which one of two transposed words was substituted with a syntactically illegal word matched for length (e.g., *He the jacket glass there*). The authors suggested that this effect is potentially indicative of sentence-level representations receiving more activation in the transposed-word sequences compared to the control sequences since there is a viable sentence representation for the former but not the latter sequence. These findings suggest that word transpositions influence performance on a variety of tasks regardless of whether participants are explicitly instructed to search for an ungrammaticality or not. Hence, it is important to consider how word transpositions may affect processing beyond the frequently employed speeded GDT (e.g., Mirault et al., 2018) and in more natural reading settings.

There have been three studies to date that have investigated how word transpositions may affect reading in alphabetic spaced languages via measuring eye movements. Since the current thesis is also focused on how transposed words affect eye movements, these three studies will be discussed in depth next.

The first study to look at word transpositions via eye movements was conducted by Rayner, Angele et al. (2013) who manipulated the parafoveal preview of two target words via the boundary paradigm (Rayner, 1975). Rayner, Angele et al. used three types of parafoveal previews. First, for some sentences the two target words remained the same and were presented in order both before and after the eyes crossed the invisible boundary – an identity preview (e.g., *My neighbor painted the white walls black.* -> *My neighbor painted the white walls black*). Secondly, the two target wordscould appear out of order before the eyes crossed the boundary, but appear in their correct order after the eyes crossed the boundary (e.g., *My neighbor painted the walls white black* -> *My neighbor painted the white walls black*). Finally, they could be substituted by two unrelated words in preview (e.g., *My neighbor painted the vodka clubs black* -> *My neighbor painted the white walls black*). Crucially, the transposed-word previews in this study did not constitute a grammaticality violation in contrast to the other studies on word transpositions discussed so far and later in the thesis. This is important because even if one were to assume that readers may be sensitive to ungrammaticality in the parafovea, given that the transposition previews in the current study did not result in an ungrammaticality, there would be no parafoveal cue to a transposition. However, when the two target words were substituted by unrelated words in preview, they did not fit with preceding sentential context. To be clear, the substitute preview words did not produce an ungrammaticality, but they were semantically anomalous in respect of sentential context. If readers were sensitive to semantic anomalies in the parafovea, then the only case in which there would be a semantic anomaly in parafoveal vision would be when the two target words were substituted. Thus, if readers can lexically identify multiple words in parallel, and further, if they are also able to recognise that an upcoming word is semantically anomalous given previous context, the substitute preview may result in a semantic parafoveal-on-foveal effect at the point of fixating the pre-target word. Conversely, since the transposed word preview is perfectly grammatical and semantically plausible, there should be no disruption on the pre-target word associated with either a transposed word or an identity preview. Contrary to these assumptions, the authors found no significant disruption on the pre-target word for FFD, GD and GPT (i.e., prior to the eyes crossing the boundary) for either of the two invalid preview conditions.

Rayner, Angele and colleagues also analysed the eye movement patterns on the target words. The authors conducted analyses on each of the two words separately. Furthermore, they conducted a second set of analyses considering both target words as a single region of interest. The disruption associated with a transposed or substituted preview was comparable on early measures such as FFD and GD for the combined region and the first target word. However, the disruption to processing reflected in later measures such as GPT was increased for the substitute preview compared to the transposed-word preview on both the combined target region and the first target word.

Moreover, the gaze duration on the first target word was increased for the substitute compared to the transposed preview conditions specifically when the pre-target word had not been skipped during first-pass reading. Overall, the authors suggested that changes to word order in preview result in significant disruption to reading once the invisible boundary is crossed. As to why the substitute preview caused more disruption to processing than the transposed-word preview, the authors suggested that the degree of overlap between the target words and their transposed-word preview was higher than the overlap between the target words and their substitute word preview. That is, readers were potentially able to extract some low level sublexical information from a transposed word preview which was naturally not possible for a substitute word preview.

Mirault et al. (2020) on the other hand, investigated whether word transpositions disrupt local eye movement measures in a virtual reality eye-tracking experiment investigating reading. The authors used the same materials as Mirault et al. (2018) and the critical comparison was between two ungrammatical sequences. These were sentences containing a word transposition only (e.g., *The white was cat big*) and sentences containing both a word transposition and a syntactically illegal final word (e.g., *The white was cat slowly*). Additionally, participants were required to make a grammaticality judgment for each sequence. Participants were instructed to make speeded grammaticality decisions, however there was no time out/limit set for each trial, meaning that each participant could technically take as long as they needed to reach a decision. The authors obtained the standard *Transposed-Word* effect on accuracy both when the two target words were fixated in their order of presentation, as well as when they were fixated out of order. Moreover, total sentence reading times were inflated for sentences containing a word transposition only versus sentences containing both a word transposition and a syntactically illegal final word. However, they found no evidence of disruption to the total viewing times, skipping and refixation probabilities on either of the two critical transposed words alone or when investigated as a single transposed word pair region. The authors interpreted this finding as evidence in favour of parallel processing such that word order violations do not disrupt local oculomotor behaviour but instead the disruption is only visible in global measures such as RTs and total sentence reading times.

An alternative explanation is possible, however. Since both sequences included a word transposition, it is possible that the lack of effects on local eye movement measures simply indicated that readers were not sensitive to the syntactic legality of the final word at the point of fixating either of the target words. That is to say, in order for the syntactic illegality of the sentence final word to affect reading times at the transposed word pair, it would have to be processed (at least to the degree that its syntactic class was available) at a point in the sentence at least one or two words prior to it. Such a possibility may be considered unlikely, and from this perspective, therefore, the similarity in local reading times at the transposed word pair is perhaps unsurprising. This alternative interpretation does not necessitate parallel processing and furthermore is perfectly compatible with a serial processing framework such as E-Z Reader (Reichle, 2011). Furthermore, since the syntactically illegal final word was only present in conjunction with a word transposition but not on its own, it is difficult to know whether having two violations to syntax (i.e., a word transposition and a syntactically illegal final word) versus one violation (i.e., a word transposition only) differentially influenced the observed effects. In other words, it is unclear whether word transpositions are special or whether it is a simple case of how easy it is to judge a sequence of five words as ungrammatical when 40% of the words are incorrect versus when 60% of the words are incorrect.

In contrast to Mirault et al. (2020), Huang and Staub (2021a) focused on the comparison between grammatically correct sentences and sentences containing a word transposition only. Furthermore, the authors factorially manipulated word length (both short versus one long one short) and word class (both open versus first open second closed versus first closed second open) to determine whether these variables might modulate the effects of word transpositions. Nouns, adjectives, and verbs were used as the open class words while prepositions, determiners, and pronouns were used as closed class words. For instance, readers were presented with sentences such as *She repeated favorite her song ten times* in which the transposition is created by one long open class word (*favorite*) and one short closed class word (*her*) or *The fragile cup red shattered into pieces* where the two transposed words are short and open class (*cup red*) or *The boy on sat the school bus* in which the transposition is created by two short words but one word is closed class (*on*) and the other word is open class (*boy*). In addition to transposed-word and grammatically correct sentences, readers were also presented with incomplete sentences which acted as fillers in the experiment (e.g., *I see the incredible talent in your*). This was done to encourage participants to read each sentence to the end. The filler sentences were not included in the analyses. Furthermore, unlike the other two studies, Huang and Staub chose to use Bayesian mixed effects modelling. Consequently, their results may not be interpreted in terms of significant or non-significant effects. Instead, their results may show how likely it is that there is no difference between two or more conditions for a specific measure.

There are three aspects of their findings of note. Firstly, Huang and Staub (2021a: Experiment 1) found a *Transposed-Word* effect on accuracy (they did not investigate RTs). This is important because given the design of the experiment, OB1 Reader (Snell, van Leipsig, et al., 2018) would predict that word transpositions of two words that are substantially different in length should have little to no influence on accuracy. This is because word length cues are a primary determinant of what position each word is assigned. For example, if the transposition is created by an eight-letter word (*favorite*) and a short word (*her*) in the third and fourth positions in a sentence, the sentence-level representation should search for a viable eight-letter word in position three and a viable three-letter word in position four. OB1 Reader does state that syntactic expectations in principle can explain how such word transpositions between two words of vastly different lengths may not be detected. However, again, using the example of *She repeated favorite her song ten times* the initial preference for word position coding should be for an eight-letter word in position three and a three-letter word in position four. Hence, even if there is a strong syntactic expectation for an adjective or a noun at position four instead of a pronoun, the sentence-level representation should preferentially assign a word belonging to an expected syntactic category that is comparable in length (e.g., two, three or four-letters long) to position four in the sentence. Therefore, it stands to reason that the word *song* is a more likely candidate for position four than the word *favorite*.

Secondly, the authors found that accuracy was modulated by whether either of the two transposed words was directly fixated during first-pass reading or not. Again, this finding is partially inconsistent with OB1 Reader since fixation patterns should not change regardless of the failure or success in detecting a transposition. This is because lexical and post-lexical information integration is handled by a sentence-level representation in OB1 Reader and functions independently from oculomotor control mechanisms. Therefore, the mechanisms which determine where the eyes look and how long a word is viewed should not affect or be affected by the mechanisms which assign words to their respective positions within a sentence.

Thirdly, the authors found significant and early oculomotor disruption associated with the presence of a transposition. This disruption was evident in longer reading times on the second word of the transposed word pair since the sentences became ungrammatical at that point. Reading times were inflated for all four reading time measures examined (i.e., FFD; GD; GPT and TVT). Furthermore, this disruption was only present on trials during which the participants overtly judged the sentence as ungrammatical but not on trials where they made the incorrect grammaticality decision. Similarly to the previous two points, this is problematic for OB1 Reader since the model postulates word transpositions should not disrupt local eye movement measures due to the implementation of a sentence-level representation that deals with word position coding and is not dependant on reading times.

The authors were clearly critical of the OB1 account of *Transposed-Word* effects and instead suggested that their findings could be reconciled with the serial account of E-Z Reader (Reichle, 2011). According to the baseline E-Z Reader model (Reichle, Warren, et al., 2009), integration of word *n* needs to be completed before integration of word *n+1* into context can begin. Such a mechanism would preclude readers from failing to detect word transpositions unless they have viewed, and therefore identified, the words out of order and should result in significant and early disruption to reading times at the point of encountering a word transposition that causes an ungrammaticality. Instead, Huang and Staub (2021a, b) propose that the integration of two consecutive words into context can overlap. This suggestion does not directly reconcile serial processing accounts with the lack of disruption associated with the word transposition when participants failed to detect it. However, such overlap in integration allows for rational inference to determine the order in which words are integrated into context. In other words, while words may be identified serially and sequentially, they may not be exclusively integrated in the order in which they are identified due to the rapid and immediate development of higher-order post-lexical expectations.

1.9. Summary and outline of the present thesis

In summary, eye-tracking reading research has shown a sensitivity to a variety of visual, orthographic, phonological, semantic, syntactic, pragmatic, and discourse violations in foveal vision both in terms of early and late reading times measures (e.g., FFD; TVT) as well as saccade targeting measures (e.g., SP; ROut). The extent of sensitivity to semantic and syntactic violations in the parafovea is, on the other hand, debated (see Rayner, 2009; Schotter et al., 2012 for reviews). Additionally, multiple advanced experimental paradigms and computational models of oculomotor control in reading have been developed over the past half a century to account for the complex empirical findings in the field and to drive the development of further hypotheses. In the present thesis, I will present three eye-tracking experiments that investigate the recently established *Transposed-Word* effect during natural reading.

The first experiment (Chapter 2) aims to disentangle the influences of word transpositions and final word syntactic legality on both reading and grammaticality decisions. The experiment employs a fully factorial design to independently manipulate the syntactic legality of the final word in the sentence, and the word order of the third and fourth sentence words. Previous studies on the *Transposed-Word* effect (Hossain & White, 2023; Liu et al., 2020, 2021, 2022; Mirault et al., 2018, 2020) have compared performance for sentences containing a word transposition only versus sentences containing a word transposition and a syntactically illegal final word on the GDT. The task requires participants to simply indicate whether a sentence is grammatical or not regardless of how many violations of syntax are included in the sentence. Hence, if participants are only indicating that a sentence is ungrammatical, performance for both types of sentences should be comparable. Given the existing findings however, Mirault et al. have argued that the lower accuracy and longer RTs for sentences with a transposition only than for sentences containing two violations is indicative of noisy word position coding as per the OB1 Reader (Snell, van Leipsig, et al., 2018). However, there is an alternative explanation for these findings such that it may simply be easier to detect an ungrammaticality when the total number of ungrammaticalities in a sentence is higher. Therefore, orthogonally manipulating the factors of word transposition and final word syntactic legality would provide novel insight into whether word *Transposed-Word* effects should be taken as evidence for parallel processing or may arise due to the more general difficulty in detecting ungrammaticality in the presence of one versus more ungrammaticalities.

The second experiment (Chapter 3) aims to provide the first eye-tracking investigation into the role of parafoveal processing concerning *Transposed-Word* effects. It was partially motivated by an earlier study by Rayner, Angele, et al. (2013) which investigated the role of word transpositions that did not constitute a grammaticality violation exclusively in parafoveal vision. The authors found via the boundary paradigm (Rayner, 1975) that reading on the critical post-boundary words was disrupted the most when in preview the words were substituted (e.g., *vodka clubs* - *white walls*), and there was significant but lesser disruption when the words were transposed in preview (*walls white* - *white walls*) in comparison to when there was no preview change (e.g., *white walls* - *white walls*). The authors suggested that even though the transposed preview was invalid, the mismatch was reduced since word *n+1* in preview became word *n+2* after the eyes crossed the boundary while in the preview substitution condition, there was no match between either of the two target words and their respective previews. The second experiment of my Ph.D. utilises the boundary paradigm (Rayner, 1975) to independently manipulate the order of two target words in the parafovea and the fovea. When the two target words are transposed, they constitute a grammaticality violation which presents a key deviation from the design of the study by Rayner, Angele, et al. (2013). This is because Experiment 2 allows me to investigate whether the presence of an ungrammaticality created by transposing two adjacent words can be detected parafoveally. Furthermore, in previous studies using the boundary paradigm (Rayner, 1975) there has been a robust preview benefit associated with presenting the same word in parafoveal and foveal vision (identity preview) when reading grammatically correct sentences (see Schotter et al., 2012). Experiment 2 allows me to investigate whether this standard preview benefit would be maintained when the target words constitute a grammaticality violation.

The third experiment (Chapter 4) uses the boundary paradigm as well, to compare the effects of word transpositions to letter masks on the first, second, or both critical target words in parafoveal vision. This experiment is aimed at investigating the extent of parafoveal processing beyond the upcoming word *n+1* in parafoveal vision. Furthermore, since any ungrammaticality is only presented parafoveally but not foveally, in Experiment 3 participants are not asked to make any grammaticality decisions and instead are tasked with answering comprehension questions on a proportion of trials. This is an important distinction from the first two experiments because it provides a further test of the effects of parafoveally presented word transpositions in more natural reading conditions. While the existence of parafoveal-on-foveal and parafoveal preview n+2 effects is debated (see Vasilev & Angele, 2017; section 1.5.2), none of the existing research has utilised ungrammatical word transpositions as a parafoveal preview manipulation. Hence, comparing parafoveal word transpositions to letter-string masks would enable me to separate any potential differences in eye movement patterns that are driven by the processing of parafoveally available orthographic information from differences driven by the processing of the grammaticality or lexical characteristics of the upcoming words in the parafovea. If readers do indeed extract lexical and/ or syntactic information parafoveally then the transposed-word preview should produce a pattern of eye movements that differs from the ones produced by the letter-mask preview.

All three experiments of my Ph.D. serve to expand on current knowledge of how word transpositions affect cognitive processing under more ecologically valid natural reading conditions than previous research that has used the speeded GDT (e.g., Mirault et al., 2018). Furthermore, all three experiments present an empirical test of existing computational models of oculomotor control during reading and are aimed to provide novel insight into the ongoing parallel versus serial lexical processing debate in the eye movement reading literature (see Snell & Grainger, 2019a; Zang, 2019 for discussions).

In the final discussion chapter, I will consider the implications of the findings of each experiment of my Ph.D. concerning the theoretical accounts that have been put forward to explain the *Transposed-Word* effect in the context of existing empirical research on the topic.

I would like to note that both Experiment 2, and Experiment 3 were designed and conducted as co-registered eye-tracking and EEG experiments. This was done in order to examine the neural correlates of *Transposed-Word* effects in the eye movement record. As a consequence, for both Experiments 2 and 3 reported here, I present eye movement data which is matched with the electroencephalography data cleaned and pruned following the procedures outlined by Degno et al. (2021).

My Ph.D. was significantly and adversely affected by the COVID 19 pandemic. The closure of all laboratories and campus facilities during the pandemic meant I could only start running the experiments of my Ph.D. a year and a half after my period of registered study started. Consequently, while I have successfully finished data collection for all three experiments, I have not been able to complete the analyses for the EEG data I obtained. These analyses are computationally taxing and complex meaning that they require a lot of time to run, verify and interpret. It is for these reasons that in this thesis I only present the eye movement data and findings Experiments 2 and 3 of my Ph.D.

I remain fully committed to writing up both Experiments 2 and 3 as co-registration studies (reporting both the eye movement and EEG results in full) for publication in peer-reviewed scientific journals in the period immediately after the submission of my thesis. I note also that Experiment 1 of this thesis was an eye movement study alone and this experiment has been written up and submitted for publication.

Chapter 2. Word order effects in English sentence reading

2.1. Introduction

There exists a large body of evidence suggesting that processing of letters within a word occurs in a parallel fashion, and the encoding of letter positions is subject to a level of uncertainty (e.g., Gomez et al., 2008). This body of evidence is, to some extent, based on letter transposition effects, whereby nonwords created by swapping the order of two letters within the original word (e.g., *dcotor* from *doctor*) are responded to, or read faster, compared to nonwords created by substituting two of the letters within the original word (e.g., *dmator*; see Johnson et al., 2007; Pagan et al., 2021).

Higher order transpositions at the level of morphemes and words have remained relatively under investigated with only a single study on word transpositions prior to 2018 (i.e., Rayner, Angele, et al., 2013) in English. Rayner, Angele, and colleagues investigated the influence of word transpositions in preview via the boundary paradigm (Rayner, 1975; see Section 1.5.2) and compared reading times when participants received an identical preview, a transposed-word preview, or a substituted word preview. The transposed-word preview did not constitute a grammaticality violation at the point of first crossing the invisible boundary. Their findings suggest that both invalid preview conditions caused disruption with longer reading times on the post-boundary word. However, that cost was graded such that when the preview consisted of substitute words, reading times were longer than when the preview consisted of a transposed word pair.

2.1.1. The Transposed-Word effect

Work on word transpositions in the past six years has been focused primarily on the detectability of transposed words that create a grammaticality violation (e.g., Mirault et al., 2018). This line of research has prompted experimentalists to consider whether positional encoding and processing of words within sentences or word strings might be similar to positional encoding and processing of letters (Dufour et al., 2022; Huang & Staub 2021a, 2021b, 2022, 2023; Hossain & White, 2023; Liu et al., 2020, 2021, 2022, 2024; Milledge et al., 2023; Mirault et al., 2018, 2020; Mirault, DeClerk, et al., 2022, 2023; Mirault, Leflaec et al., 2022; Mirault, Vandendaele, et al., 2022, 2023; Pegado & Grainger 2019, 2020, 2021; Snell & Grainger, 2019a, 2019b; Snell & Melo, 2024; Snell, Mirault, et al., 2023; Spinelli et al., 2024; Tiffin-Richards, 2024; Wen et al., 2019, 2021a, 2021b, 2022, 2024).

A striking finding from Mirault et al. (2018, 2020) is that readers sometimes fail to notice word transpositions within a sentence (e.g., *The white was cat big*) if a plausible representation of that sentence can be formed (e.g., when the words can be rearranged to be meaningful and syntactically correct, as in *The white cat was big*). This failure to detect a word transposition is particularly pronounced when compared to a situation where such a representation is not possible (e.g., when the final word in a string precludes syntactic legality of that string, as in *The white was cat slowly*). The *Transposed-Word* effect has also been found to be greater (i.e., slower responses and more detection failures) when the transposed word pair is internal (e.g., *The can man run*) compared to external (e.g., *Run man can the*; Snell & Grainger, 2019b).

The *Transposed-Word* effect has been replicated in different languages (in Chinese: Liu et al., 2020, 2021, 2022; Dutch: Snell & Melo, 2024; German: Tiffin-Richards, 2024; and English: Hossain & White, 2023; Huang & Staub, 2021a, 2022, 2023; Milledge et al., 2023). Effects of word transpositions with comparable magnitudes have been observed in both speeded (e.g., Liu et al., 2020, 2021; Mirault et al., 2018) and unspeeded (e.g., Liu et al., 2021) grammaticality decisions suggesting that the effect is robust even when participants did not need to sacrifice accuracy to perform the task as quickly as possible (a speed-accuracy trade-off). Beyond grammaticality decisions, word transpositions have been shown to affect change detection (e.g., Pegado & Grainger, 2020) such that participants are worse at detecting a change that was created by transposing two words than by replacing them, suggesting that word transpositions may be difficult to detect regardless of the specific task constraints.

2.1.2. Eye-movement control models and the Transposed-Word effect

The recently established OB1-Reader model (Snell, van Leipsig, et al., 2018; see Section 1.6.3) directly aims to explain *Transposed-Word* effects. The model specifies that multiple words are lexically processed in parallel during natural reading. According to OB1-Reader, semantic information is not immediately integrated across words during natural reading (Snell & Grainger, 2019a). Instead, when words are identified in parallel, semantic information associated with each of those words is tied to the particular location assigned to that word in the sentence frame via a spatiotopic sentence-level representation. The individual semantic meanings of the words are then integrated with sentential context at a later, integration, stage of processing. In this way, the model allows for parallel lexical identification of words whilst clearly specifying that the semantic characteristics of words that appear later in the sentence will not influence processing of words that appear earlier in the sentence.

According to Snell, van Leipsig, et al. (2018), readers form a sentence-level spatiotopic representation upon first fixating a sentence, based on word length and syntactic expectations. Based on these assumptions, then, it is quite possible that word order determined by the reader might not match the actual order of the words that appear within a sentence. In this way, the transposed word phenomenon, that is, failure to detect transposed words in otherwise grammatically correct sentences, occurs due to successful location tagging of words in relation to their correct positions in a sentence. Furthermore, since word position coding is handled by the spatiotopic sentence-level representation in working memory, any disruption associated with detecting a word transposition (i.e., detecting the ungrammaticality) should be observable in global measures such as RTs and accuracy on the

GDT. Conversely, no disruption should be observable in the local eye-movement measures on the critical transposed words.

The SWIFT (Engbert et al., 2002; 2005; see Section 1.6.2) is a parallel gradient model that preceded OB1 Reader. According to SWIFT, multiple words are lexically processed in parallel during sentence reading. Because lexical identification of multiple words can occur in parallel, the semantic and lexical characteristics of the upcoming two (or so) words can affect processing of the currently fixated word, thus, potentially causing semantic and/or syntactic parafoveal-on-foveal effects. Thus, SWIFT stipulates that readers process words based on the order in which they are identified rather than the order in which they appear within a sentence.

The SWIFT model may offer an account of how transposed words that appear within a sentence might be processed successfully, although this explanation would require that out of order lexical identification should occur for each occasion that readers failed to detect a transposition. Although this is possible, it is important to understand that according to the SWIFT model, parafoveal words further from the point of fixation most often receive less activation than parafoveal words closer to fixation due to visual constraints. This means that words further into the parafovea are usually identified more slowly than those closer, and therefore, an explanation based on consistent out of order lexical identification seems unlikely. The predictions based on the SWIFT model are also consistent with the predictions of its recent successor model, SEAM (Rabe et al., 2024), which largely incorporates the lexical processing assumptions of SWIFT. Given this, for the remainder of the chapter, SWIFT and SEAM will be considered together.

In contrast to parallel models, the E-Z Reader Model (Reichle et al., 1998; Reichle, 2011; see Section 1.6.1) is a Serial Attention Shift model. According to this model, words are processed in stages, serially and sequentially, whilst a process of integration runs in the background. Recognition of a word during reading occurs initially via a preattentive visual stage of processing whereby the visual characteristics of the upcoming word are processed. This is followed by an early lexical stage of processing (a familiarity check) wherein readers discern the letter make-up, that is, the orthography and potentially the phonology of the word. Next, during a later lexical stage, full lexical access of the word occurs, and this takes place as the word is directly fixated. Most of the time, the word to the right of the currently fixated word, in the parafovea, can only be pre-processed up to and including the early lexical stage before it is fixated.

According to the stipulations of E-Z Reader, therefore, readers should only process words in the order they are viewed, and consequently, the only way in which E-Z Reader can explain why readers sometimes fail to detect transposed words is by readers fixating those words out of order. Like SEAM with SWIFT, a successor model for E-Z Reader has recently been put forward, namely the Über Reader (Reichle, 2021; Veldre et al., 2020). Again, most of the assumptions associated with lexical processing in relation to the formation of oculomotor commitments (where and when to move the eyes) in E-Z Reader have been carried over to Über Reader, and therefore, hereafter I will consider E-Z Reader and Über Reader together.

2.1.3. Alternative explanations of the Transposed-Word effect

As noted earlier, the account of processing offered by the OB1-Reader model has been used to explain *Transposed-Word* effects (e.g., Mirault et al., 2018, 2020). However, some research has challenged this interpretation (Hossain & White, 2023; Huang & Staub, 2021a, 2022, 2023; Liu et al., 2020, 2021, 2022; Milledge et al., 2023; though see also Mirault, Vandendaele et al., 2022; Snell & Melo, 2024).

Liu et al. (2022) included word transpositions in grammatical and ungrammatical base sentences, where the final word could be syntactically legal or illegal. The authors found a word transposition effect regardless of whether the sentences were presented one word at a time in the middle of the screen, to encourage serial lexical encoding, or presented all at once on the screen, to encourage parallel processing. Although notably the *Transposed-Word* effect in the serial mode presentation was weaker and only reliably obtained for accuracy and not for RTs while in the parallel mode, the effect was observed for both accuracy and RTs. This led Liu et al. (2022) to conclude that transposition effects do not necessarily indicate that word processing needs to occur in parallel (see also Huang & Staub, 2023). Instead, they argued that the failure to notice words out of order might be explained in line with *noisy-channel* theories (e.g., Gibson et al., 2013). That is, readers might initially detect the word transposition, and thus, the sentence ungrammaticality, but due to top-down predictive mechanisms, such as syntactic and semantic expectations, they might still be able to form a plausible interpretation and adopt this, thereby overriding the syntactic violation (see also Liu et al., 2020). Subsequently, Huang and Staub (2023) found a similar pattern of effects when comparing serial versus parallel mode presentation and argued that the presence of a *Transposed-Word* effect in serial presentation mode is clear evidence against parallel processing.

Mirault Vandendaele et al. (2022) observed the same pattern of effects and argued that the lack of a *Transposed-Word* effect on RTs when sentences are presented serially provides evidence in favour of the OB1 Reader account since syntactic expectations still drive word position coding even in the absence of a defined spatiotopic sentence-level map. Recently, Snell and Melo (2024) failed to find any *Transposed-Word* effects when serially presenting sentences word by word in Dutch which they argued is further evidence that word position coding is based on word length cues and syntactic expectations and these exert a weaker influence in serial presentation scenarios than in parallel (i.e., natural) presentation scenarios. Hossain and White (2023), recently provided evidence, however, that *Transposed-Word* effects can be reliably obtained for RTs and accuracy when presenting words serially when the presentation speed was matched to that of the reader. In addition, Milledge et al. (2023)

observed a robust *Transposed-Word* effect on accuracy in serial presentation mode when words were presented one at a time but in their correct sentential positions rather than in the same central location.

Whether the *noisy-channel* theories (e.g., Gibson et al., 2013) are able to explain the transposition effects was directly examined by Huang and Staub (2021a) via two eye tracking experiments. The authors hypothesised that if those theories are correct, the data should show an effect of word transposition in early eye movement measures, even when readers fail to notice that transposition, and that this effect should be reduced in later measures once readers override the grammatical violation in favour of a plausible sentence representation. In their first experiment, participants engaged in a GDT. In their second experiment, participants were asked to read each sentence and then answer a question concerning its grammaticality, or instead a comprehension question pertaining to its meaning. Participants were not forewarned of the type of question they would be required to answer on any particular trial.

Regardless of the type of questions participants were required to answer in the two experiments, Huang and Staub found that there were more grammaticality decision errors when judging sentences with transposed words compared to grammatically correct sentences, thus confirming that sometimes readers do fail to notice words out of order. Moreover, the eye movement data showed disruption (i.e., longer reading times for both early and later measures) at the point of ungrammaticality (i.e., the second word of the transposed pair) only when participants detected the violation, but not when they failed to detect that the words were out of order. Early reading times were comparable at the point of ungrammaticality between grammatically correct sentences and sentences with a transposition that was not noticed. Note, however, for this latter condition, the eye movement record did show some disruption to reading on later measures. Specifically, Huang and Staub showed inflated rates of regressions to the text preceding the ungrammaticality and increased total viewing times for the word preceding the ungrammaticality for Experiment 1, and increased regression-out rates for the word preceding the ungrammaticality for Experiment 2. These findings were interpreted by the authors as evidence that some form of top-down mechanism must play a role in word processing, as readers do, in fact, sometimes fail to notice the word transpositions. However, the authors argued that any such top-down mechanism must operate much earlier than *noisy-channel* accounts would predict, as there were no effects of word transposition in the early eye movement measures when readers failed to notice the violation. Thus, Huang and Staub (2021a) proposed that word transpositions can be explained by a serial processing account (e.g., the E-Z Reader model, Reichle et al., 1998; Reichle, 2011) in which semantic and syntactic integration can proceed in a non-incremental fashion, with integration of word *n+1* at times taking place before integration of word *n* is completed (see Huang & Staub, 2021b for a detailed discussion).

Although more research is needed to establish what causes readers to fail to notice word transpositions, two additional findings from Huang and Staub (2021a) seem to contradict the OB1-Reader predictions for this effect. In their two experiments, Huang and Staub transposed target word pairs in the third and fourth position of the sentence, such that the target words were either both short, or one of them was long (short: 2-4 letters, long: 8-11 letters). According to OB1-Reader, for a word to assume a particular sentential position, it must match the perceived length of the word expected in that position (based on parafoveal and peripheral processing). If this is correct, then it follows that when transposed words are of different lengths, particularly when that length difference is substantial, readers should detect the transposition. However, in the Huang and Staub (2021a) experiments, readers failed to detect a transposition even when the target words were of very different length. It is difficult to understand why this would be the case given the stipulations of the OB1-Reader model. Additionally, in Experiment 1, Huang and Staub found that the transposition effect on accuracy was modulated by the likelihood of skipping either of the words in the transposed word pair. When participants skipped either of the transposed words, they were more likely to judge the sentence as grammatically correct. This is problematic for the OB1 Reader model because it shows that changes in local eye movement behaviour in and around transposed words modulate ungrammaticality detection. However, according to the OB1 Reader model, this should not be the case as word transpositions should be handled by the spatiotopic sentence-level mechanism and should not influence or be influenced by local eye-movement behaviour.

Similar issues are present in Mirault et al.'s (2020) study, where participants' eye movements were recorded via a virtual reality setup. In their experiment, the authors used stimuli from Mirault et al. (2018), such that a transposed target word pair was embedded in a sentence with either a syntactically legal, or an illegal, final word. Eye movements were recorded and compared for sentences containing a transposition and an illegal final word against those for sentences with a transposition and a legal final word.

Mirault et al. (2020) replicated the *Transposed-Word* effect (lower accuracy for sentences with a transposition and a legal final word than for sentences with a transposition and an illegal final word). However, the authors did not observe any effect of word transposition on the eye movement behaviour on the target word pair. They did, however, find longer total sentence reading times for the sentences with a transposition and a legal final word compared to the sentences with a transposition and an illegal final word. At first glance these results might seem to support the OB1-Reader model, in that readers failed to notice the transposition both in terms of accuracy and eye movement patterns (at least in the local measures). However, Mirault et al. (2018, 2020) did not match the two target words for length (for 38% of trials the transposed words differed in length by more than one letter), suggesting that the observed effects could not be attributed to the successful implementation of a spatiotopic sentence-level mechanism as per OB1 Reader (Snell, van Leipsig, et al., 2018). Furthermore, the average lexical frequency across the five words comprising the sentence was high and was not controlled, and it is unclear to what extent each target was predictable given prior context. In addition, it is likely that no differences were observed in the local eye movement measures for the two targets as the authors compared reading times for transposed words in sentences with a legal versus illegal final word. That is to say, the lack of a difference was likely due to the fact that the reading times were compared for word pairs that were transposed in both experimental conditions.

2.2. The present experiment

The present experiment aimed to expand on previous research on word transposition effects whilst considering the existing computational models of eye movement control during reading, namely E-Z Reader (Reichle et al., 1998; Reichle, 2011) and Über Reader (Reichle, 2021; Veldre et al., 2020), SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024), and OB1-Reader (Snell, van Leipsig, et al., 2018). Further, since Mirault et al. (2018; 2020) only directly compared sentences containing a transposition with a syntactically legal versus illegal final word, I sought to disentangle effects arising due to word transpositions and effects arising due to the syntactic legality of the final word of the sentence. This is important because the GDT is aimed at exploring how well participants are able to detect ungrammaticality in sentences. Note, though, that the task does not require participants to quantify how many ungrammaticalities they have detected, just whether they have, or have not, detected one. It is possible that the magnitude of *Transposed-Word* effects may be increased because having two ungrammaticalities in a sentence facilitates ungrammaticality detection, while a single ungrammaticality may be less easy to detect. Hence, adding a new condition in the current experiment, in which readers are presented with sentences containing a single violation to syntax that is not created by transposing two adjacent words, might allow me to determine whether the magnitude of the *Transposed-Word* effect is decreased for sentences containing two, versus one, violation to syntax.

I, therefore, manipulated final word grammaticality independently from word transposition, such that the final word could be grammatical or ungrammatical, and the third and fourth word of each sentence were presented in their correct order or transposed. I also ensured that the first point of ungrammaticality was always the third word of each sentence (i.e., the first word of the transposed pair). Furthermore, since in Mirault et al. (2018; 2020) word frequency and length were not controlled, and in Huang and Staub (2021a) they were manipulated, I chose to match the target words on both characteristics. Matching target words on frequency and length should maximise the chances of participants failing to detect the word transposition, and thus allow me to examine whether, even under these circumstances, eye movements might show some disruption.

Condition	Sentence
Final grammatical word & Non-Transposed targets	Their mother felt very pleased.
Final grammatical word & Transposed targets	Their mother very felt pleased.
Final ungrammatical word & Non-Transposed targets	Their mother felt very grandma.
Final ungrammatical word & Transposed targets	Their mother very felt grandma.

Figure 2.1. An example of the stimuli used in Experiment 1.

The five regions of interest are indicated as follows: (1) The First Word is presented with a single underline; (2) the Pre-Target Word is presented with a dashed underline; (3) Target Word 1 and Target Word 2 are presented with a double underline; (4) the Post-Target Word is presented with a dotted underline. None of the words and sentences were underlined within the actual experiment.

Based on Mirault et al. (2018, 2020) and Huang and Staub (2021a Experiment 1), I predicted that readers would be significantly worse at judging the grammaticality of sentences containing only a transposition compared to sentences with a transposition and an ungrammatical final word. I also anticipated that the proportion of correct responses would be reduced for sentences containing only a word transposition than for correct sentences, though of course, participants would make *no* responses for the former and *yes* responses for the latter. I had two diverging expectations regarding performance on sentences containing only a syntactically illegal final word in comparison to sentences containing two violations to syntax. First, if the *Transposed-Word* effect is driven by the ease with which an ungrammaticality is detected in the presence of two compared to one syntactic violation, then participants should be better (i.e., more accurate and faster) in detecting the ungrammaticality in sentences with a word transposition and a final syntactically illegal word compared with sentences containing a final syntactically illegal word only. Although there is to date no direct evidence from GDT studies for written text, these predictions are supported by evidence from studies in the auditory domain (e.g., Devescovi et al., 1997; Lu et al., 2000). If, on the other hand, the OB1 Reader account of processing is correct, then performance for sentences containing a final ungrammatical word only, should be comparable to performance for sentences containing two syntactic violations since no amount of rearrangement of word order would produce a grammatical sentence in either case. In line with this reasoning, performance for sentences containing a final ungrammatical word only should be elevated in comparison to sentences containing a transposition only, since the former cannot be resolved into a grammatical sentence, while the latter, can.

With respect to eye movements, my predictions were driven both by existing research on *Transposed-Word* effects, as well as by the different computational accounts discussed in section 2.1.2. According to E-Z Reader (Reichle, 2011) and Über Reader (Reichle, 2021; Veldre et al., 2020), there should be a significant effect of transposition starting at Target Word 1 when the two target words were presented out of order, with longer first-pass reading times when words were transposed than when they were presented in the correct order. E-Z Reader (and Über Reader) would also predict a significant effect of the grammatical legality of the final word at that word (i.e., post-target), as a serial perspective does not assume parafoveal processing of semantic-syntactic information. I might also expect to observe a significant interaction between transposition and final word grammaticality with a larger transposition effect when the final word was ungrammatical compared to grammatical. I make this prediction based on Mirault et al. (2020), who showed that decision times were shorter for sentences containing two than one ungrammaticality.

If readers process words in parallel, as per SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024), there should be an effect of transposition, such that disruption should occur as early as on the First Word of the sentence (in line with a processing span of two words to the right of the fixated word). I should also find an effect of final word grammaticality as early as fixations on the third word (Target Word 1) in the sentence, with longer reading times when the final word was ungrammatical than when it was grammatical. Finally with respect to SWIFT and SEAM, there should also be an interaction between final word grammaticality and transposition, as early as fixations on the third region of interest (Target Word 1) in the sentence, with a larger effect of transposition when the final word was ungrammatical compared to grammatical.

If readers process words in parallel, as per the OB1 Reader model (Snell, van Leipsig, et al., 2018), word transpositions should have no effect on early eye movement measures on the transposed word pair (i.e., the Target Word 1 and Target Word 2 regions). OB1 would predict an effect of final word grammaticality, but this should only occur on the fifth word (post-target) in the sentence. Reading times should be longer for ungrammatical than grammatical final words. Finally, there should be no interactive effects for reading time measures on any of the five words given that the transposition should be handled by the sentence-level spatiotopic mechanism, and therefore, processing of individual words should be unaffected.

2.3. Methods

2.3.1. Power Analysis

To determine the target sample size for the present experiment, I conducted a power analysis via the PANGEA software by Westfall (2016; http://jakewestfall.org/pangea/) based on Mirault et al. (2020) and Cohen's d_z of 0.392 for sentence total viewing times. The analysis indicated that with an effect of this size and 108 stimuli, 40 participants would be needed to obtain a power ≥ 0.9.

2.3.2. Participants

I tested a total of 45 native English speakers with no known reading impairments and normal or corrected-to-normal vision from the student and staff community of the University of Central Lancashire. Participants were recruited via SONA and social media posts to participate in the eye-tracking experiment. All participants received £6 in Amazon vouchers or 10 course credits to take part. Five participants were excluded from data analysis (four due to technical issues, and one due to an accuracy rate lower than 80%), and the data of 40 participants (*M* = 20.8 years, *SD* = 2.5, Female = 34; Age Range = 18-30 years) was included in the final analyses.

2.3.3. Design

The study was conducted as a 2(transposition: not transposed vs. transposed) × 2(final word grammaticality: grammatical vs. ungrammatical) repeated-measures Latin-square design. With this design, each participant saw only one of the four versions of each sentence, and each sentence appeared an equal number of times across participants and its four versions.

2.3.4. Apparatus

Viewing was binocular but only participants' right eye movements were recorded using an SR Research Eyelink 1000 Plus system with a sampling rate of 1000 Hz. Participants were seated 70cm from of an LCD monitor with 1920 by 1080 FHD resolution and 240Hz refresh rate. Stimuli were presented with a horizontal offset of 240 pixels from the centre of the monitor, in black on a grey background, and written in monospaced Courier font size 24, with 2.3 letters subtending 1° of visual angle. The experiment was designed and presented via Experiment Builder v2.3.38 (SR Research).

2.3.5. Materials

One hundred and forty sentences were initially created for the study. Each sentence was comprised of five words which constituted the five regions of interest (ROIs) used for statistical analyses: first word; pre-target word (always the second word in the sentence); target word 1 (in the third or fourth sentence position dependent on the experimental condition); target word 2 (in the fourth or third sentence position dependent on the experimental condition) and post-target word (always in the fifth and final sentence position). The order of the two target words (target word 1 and target word 2) was manipulated such that they could be presented in the non-transposed, correct order, or have their order swapped, such that the sentence contained a word transposition and became ungrammatical at the third word position (see Figure 1). Additionally, the final word (post-target) was chosen such that it could either fit the sentential context created by the first four words (final grammatical word) or be an ungrammatical continuation (final ungrammatical word).

To maximise the likelihood of obtaining at least one fixation on each word of the sentence, the pre-target, target word 1; target word 2 and post-target were at minimum four letters long. The Pre-Target word was between 4 and 7 letters long (*M* = 5.45, *SD* = 0.91) and always a different length than the target words[1]. In addition, I matched the target words for length (*M* = 4.53, *SD* = 0.72), each being between 4 and 6 letters long. Under these circumstances, according to the OB1-model (Snell, van Leipsig, et al., 2018), both the

---

[1] In one sentence out of the 108 sentences, due to human error, the pre-target word had the same length as the two target words. For this reason, data from trials with this stimulus was not included in the analyses.

transposed words (target word 1 and target word 2) should be very likely to be assigned to their syntactically correct positions, thereby maximising the possibility that readers would fail to detect the transposition, and thus, be more likely to judge these sentences as being grammatically correct. The grammatically correct and incorrect post-target words were also matched for length (*M* = 5.86, *SD* = 1.52), each being between 4 and 9 letters long[2].

According to OB1 Reader model simulations (Snell, van Leipsig, et al., 2018), a word with higher lexical frequency in the parafovea should inhibit processing of the currently fixated word. Additionally, lexical frequency is known to have a robust effect on oculomotor behaviour as early as the first fixation during first-pass reading on a word (e.g., Rayner, 2009). Therefore, I controlled for the potential confounding effects of lexical frequency by measuring the Log Zipf lexical frequency (van Heuven et al., 2014) for all words in the sentences and matching the frequencies of the two Target words and the Pre-Target word (see Table 2.1). Moreover, the post-target words were matched for lexical frequency across the final word grammaticality conditions to ensure any effects on the Post-Target region were not confounded by a frequency mismatch (see Table 2.1).

All sentences were pre-screened for naturalness to ensure that the grammatically correct sentences were viewed as natural while the sentences with one or two grammaticality violations were viewed as unnatural. The ratings were provided by 10 native English students or staff from the University of Central Lancashire who had no known reading impairments, had normal or corrected-to-normal vision, and did not take part in the eye-tracking experiment (*M* = 20.9 years, *SD* = 3.6, range = 19-30, Female = 6). The ratings were given for each sentence in each of its four versions from 1 (extremely unlikely to hear this sentence in an

---

[2] Two sentences had post-target words that differed by one letter between the grammatical and ungrammatical conditions. These two sentences were removed from all statistical analyses, in addition to another sentence that was excluded due to issues of presentation for some participants. Thus, the analyses reported here are based on 104 stimuli. Note that with this number of stimuli the power specifications detailed earlier were still obtained. For the results based on the full set of 108 stimuli see Appendix A, Supplementary tables 1-4.

everyday conversation) to 7 (extremely likely to hear this sentence in an everyday conversation).

Predictability, similarly, to lexical frequency is known to exert robust effects on reading times measures (e.g., Clifton et al., 2016). Therefore, to ensure that there was no significant difference in predictability between the two Target words, 10 additional native English students or staff from the University of Central Lancashire with no known reading impairments and normal or corrected-to-normal vision (*M* = 21.4 years, *SD* = 2.7, range = 19-28, Female = 7) were asked to take part in a close probability task. Each participant was presented with the first two or three words of each sentence and asked to continue each sentence with the first word that came to their mind. As each participant saw the first word and pre-target twice, 40 filler sentence beginnings were inserted into the questionnaire to mitigate any potential confounds.

Repeated measures t-test analyses in R (v4.3.1; R Core Team, 2023) showed that the sentences that contained a word transposition were rated as significantly more natural when the final word was grammatical (*M* =1.35; *SD*= 0.43) than when the final word was ungrammatical (*M* = 1.02; *SD* = 0.09; $t = 7.54$, $p < .001$). Additionally, the sentences without a word transposition were given higher naturalness ratings when the final word was grammatical (*M* = 5.54, *SD* = 0.60) than when the final word was ungrammatical (*M* = 1.18; *SD* = 0.35; $t = 64.67$, $p < .001$). Moreover, sentences with a final grammatical word were rated as more natural when they did not contain a transposition than when the two target words were transposed ($t = -58.25$, $p < .001$). Similarly, sentences with a final ungrammatical word were rated as more natural when they did not contain a word transposition versus when there was a word transposition as well ($t = -4.41$, $p < .001$). It is worth noting that grammatically correct sentences were rated as more natural than sentences with a word transposition only or a final ungrammatical word only ($t = -58.25$, $p < 0.001$ and $t = 64.67$, *p*

$< .001$ respectively) which in turn were deemed more natural than sentences containing both a word transposition and a final ungrammatical word ($t = -4.41$, $p < 0.001$ and $t = 7.54$, $p < .001$ respectively). These results show a clear gradation in the perceived (un)naturalness between sentences containing no violation, sentences containing one violation (i.e., word transposition or final ungrammatical word), and sentences containing two violations (i.e., word transposition and a final ungrammatical word) to syntax. Importantly, the sentences that had a final grammatical word and did not contain a word transposition (i.e., grammatically correct sentences) were rated as natural while sentences containing one grammaticality violation (i.e., a word transposition or a final ungrammatical word) and sentences containing both a word transposition and a final ungrammatical word were rated as unnatural. These results suggested that the transposition and the ungrammatical final word were clearly seen as a violation.

The repeated measures t-test on the predictability norming study data showed that the two targets were unpredictable and that there was no significant difference in their cloze probabilities (target word 1, $M = 1.7\%$; $SD = 5.1\%$; target word 2, $M = 2.7\%$; $SD = 7.1\%$; $t = -1.12$, $p = .265$) suggesting that any effect on the targets cannot be attributed to predictability.[3]

---

[3] The results reported here are for the set of 104 stimuli used for the analyses reported in the results section of this chapter. For the results from the pre-screens based on the full set of 108 stimuli see Appendix A, Supplementary table 2.

Table 2.1.

Descriptive statistics for frequency of words at all positions in the sentences. Frequency data were obtained and calculated via the Log Zipf scale (van Heuven et al., 2014).

Word	Frequency		t-test results (df = 206)				
	Range	*M (SD)*	first word t-value (p-value)	pre-target t-value (p-value)	target word 1 t-value (p-value)	target word 2 t-value (p-value)	post-target grammatical t-value (p-value)
first word	2.11 – 7.67	5.76 (1.45)					
pre-target	2.95 – 6.55	5.32 (0.69)	2.82 (0.005)				
target word 1	3.40 – 6.90	5.49 (0.71)	1.71 (0.088)	-1.78 (0.077)			
target word 2	3.51 – 7.19	5.48 (0.81)	1.73 (0.085)	-1.55 (0.122)	-0.10 (0.920)		
post-target grammatical	2.37 – 6.45	4.82 (0.81)			6.36 (<0.001)	5.89 (<0.001)	
post-target ungrammatical	2.39 – 6.35	4.63 (0.75)			8.59 (<0.001)	7.95 (<0.001)	-1.82 (0.070)

Note: These are the statistics for the final set of 104 sentences that were used for statistical analyses

2.3.6. Procedure

Participants were presented with an information sheet and a consent form upon arriving at the laboratory. Next, they were seated in front of a monitor and asked to place their chin and forehead on a headrest for head stabilisation purposes. Once the participants read the instructions, their eyes were calibrated via a 3-point horizontal calibration. This calibration was carried out after each break and whenever necessary, during the experiment.

Each trial started with a drift check 53 pixels (1 degree of visual angle) to the left of the first letter of the first word in the sentence. In the case of an error above the 0.3-degree threshold, calibration was carried out again. Following the drift check, a fixation cross was presented in the same location, and participants had to look at the cross for 500ms before the sentence would appear. If a participant moved their eyes and did not look at the cross for 500ms from the onset of its presentation, a new fixation cross was presented, or a recalibration was conducted. Participants read a block of 12 practice trials followed by 108 experimental trials separated into four blocks of 27 trials. Breaks were provided between each block and whenever the participant required.

When participants finished reading each stimulus, they had to determine whether the sentence they just read was grammatically correct or incorrect. They used a response box with the left button as the *no* answer and the right button as the *yes* answer. When the participants completed the last trial, they were presented with a *Thank you* message and a debrief form. After receiving their debrief form, participants left the laboratory, and their anonymised data were saved. The entire experimental session lasted on average between 45 and 60 minutes. The experiment was approved by the Ethics Committee at the School of Psychology and Computer Science at the University of Central Lancashire (Ethics Reference: SCIENCE 0150).

2.4. Results

2.4.1. Data analysis

Data were extracted in Data Viewer v4.1.211 (SR Research), and all analyses were conducted in R v4.3.1 (R Core Team, 2023) with the lme4 package v1.1-33 (Bates et al., 2015). Only fixations between 80 and 800ms that were not preceded, or followed by, a blink were considered for the analyses. Raw eye movement data were analysed for first fixation duration (FFD: the duration of the first fixation on a word), single fixation duration (SFD: the duration of the fixation when only one fixation was made on a word during first-pass reading), gaze duration (GD: the sum of all fixations on a word during first-pass reading before the eyes move away to another region), total viewing time (TVT: the total time spent reading the word) and refixation probability (RP: the probability to fixate a word more than once before moving the eyes to another word during first-pass reading) for each word within the sentence (i.e., First Word, Pre-Target, Target Word 1, Target Word 2, Post-Target). Go-past time (GPT: the duration of all fixations on the word and any regressive fixations on words to the left of the current word, before the eyes move to the right), skipping (SP: the probability of skipping a word during first-pass reading), and the probability to regress out of a word (ROut: the probability that the eyes move from the currently fixated word to another word to the left during first-pass reading) were analysed for all words within the sentence except First Word. Being the first region of interest within the sentence, no meaningful regressions could be made out of First Word (i.e., gaze duration and go-past time coincided). Additionally, as participants had to make a rightward saccade from the initial fixation cross location to the First Word, parafoveal processing of the First Word was likely different than for other regions of interest in the sentence, and thus the skipping rate would not have been comparable. Similarly, the probability to regress into a word (RIn: the probability of fixating a word after having fixated a word to the right) was analysed for all words within the sentence except the Post-Target, because there was no following word to regress from. I analysed fixation probability (FP: the probability to fixate a word) instead of skipping probability for the Post-Target only since there were no words to the right and thus not fixating the Post-Target would not be comparable to skipping sentence-internal words (i.e., Pre-Target, Target Word 1, or Target Word 2).

I used successive differences contrasts via the *contr.sdif* function in the *MASS* package for R (Venables & Ripley, 2002). With these contrasts I compared each successive level of each factor to the previous level of the same factor (e.g., transposed – not transposed). The *glmer* function with distribution specified as *gamma* and the link specified as *identity* was used to analyse RTs and the continuous eye movement variables (FFD, SFD, GD, GPT, TVT). Accuracy and the categorical eye movement variables (RP, SP, FP, ROut, RIn) were analysed with the *glmer* function, the distribution specified as *binomial*, and the link specified as *logit* (see also Lo & Andrews 2015; Veldre et al., 2017).

Given the assumptions of all computational models discussed so far (i.e., SWIFT: Engbert et al., 2005; SEAM: Rabe et al., 2024; OB1 Reader: Snell, van Leipsig et al., 2018; E-Z Reader: Reichle et al., 2011; and Über Reader: Reichle, 2021) early reading measures on First Word (i.e., FFD; SFD; GD and RP) and Pre-Target (i.e., FFD; SFD; GD; GPT; SP; RP, and ROut) should not be reliably affected by the grammaticality of the final word (Post-Target) due to the distance between the First Word or Pre-Target and the Post-Target. Therefore, for all early reading measures on First Word and Pre-Target, only the word transposition was included as a fixed factor in the generalised linear mixed effects model analyses. Both the word transposition and the final word grammaticality were included as fixed factors in the generalised linear mixed effects models analyses for the later reading measures on First Word and Pre Target (i.e., TVT and RIn) as well as for all reading measures on Target Word 1, Target Word 2, and Post-Target, and the global measures (i.e., accuracy and RTs). All models were equipped with a full random structure, with both random intercepts and slopes for subjects and items as per Barr et al. (2013). In the event that a model failed to converge, I systematically reduced the random structure by first trimming down the items level correlation. If a model still did not converge, I reran it excluding the interaction between word transposition and final word grammaticality where applicable (i.e., this step was not applied to the early reading measures models for First Word and Pre-Target since those models only contained one fixed factor) with the correlation included back in the random structure for items. If the model still did not converge, both the correlation and interaction were removed (where applicable). If the models still failed to converge, each random slope was subsequently removed one-by-one (first the transposition then the grammatical fit slope) until there remained just the minimal random structure for items. For models that failed to converge even then, the same trimming procedure was applied to the random structure for subjects until only the intercepts for items and subjects remained.

For both RTs and eye movement analyses, only observations from correct response trials were considered. Additionally, all analyses were adjusted via the Bonferroni Correction as per (Von der Malsburg & Angele, 2017) to mitigate for type 1 statistical errors.

2.4.2. Accuracy results

Overall, in respect of accuracy, participants performed the task well gaining accuracy scores of more than 90% in all conditions (see Table 2.2). Despite the good overall performance on the task, there were robust differences across conditions. There was an effect of transposition on accuracy, and an effect of whether the final word was, or was not, grammatical (see Table 2.3). As can be seen from the means in Table 2.2, these effects were largely driven by the interactive pattern such that grammaticality decision performance was substantially higher when sentences contained both a transposition and an ungrammatical final word than any of the other three conditions for which performance was very similar. I note that in considering these results, I am comparing conditions in which participants were required to make *yes* (i.e., grammatical) decisions with *no* (i.e., ungrammatical) decisions, and ordinarily such comparisons across decision types are not made (e.g., in lexical decision tasks). However, here I feel that such comparisons are important and necessary because they allow me to evaluate whether there are performance differences in the context of more, or less, evidence in favour of an outcome. That is to say, in making these comparisons, I can assess whether having three, two, one or no grammaticality violations within a sentence influences decision accuracy. And, as can be seen, it is clearly the case that having three grammatically illegal words in a sentence causes participants to make more accurate judgments than having two or only one ungrammatical word in a sentence. This result is interesting in that it suggests that evidence to make a *no* (i.e., ungrammatical) decision is cumulative such that when there are two cues to an ungrammatical sentence, that is, a transposition and an ungrammatical final word, then participants are better at forming an appropriate judgment than when only one of the two cues is present. At a more general level, the accuracy results suggest that transpositions may be easier to detect than was previously shown by Mirault et al. (2018; 2020), and hypothesised by the OB1 Reader (Snell, van Leipzig, et al., 2018). There are at least three reasons why responses were more accurate in the present study than was the case in the Mirault et al. (2018; 2020) studies. First, in Mirault et al., participants were required to make speeded grammaticality decisions, whereas in the present study there was no time limit in respect of decisions (accordingly, the summed total times for each region and the RTs in the current experiment indicate that total times were longer here than was the case in the Mirault et al. studies). Perhaps the requirement to make more rapid responses in the Mirault et al. studies may have led participants to make more errors as per standard speed-accuracy relations (see also Liu et al., 2021). Second, due to the fully rotated experimental design, 75% of the present stimuli required a *no* (ungrammatical) decision, whilst 25% required a *yes* (grammatical decision), thus, some of this effect may arise as a consequence of response bias. Third, the transpositions in the current study were always formed from two quite long words (4-6 letters) and most often at least one of these words was a content word. In contrast, transposed word pairs in Mirault et al. (2018; 2020) (and in Huang & Staub, 2021a) were often formed from at least one shorter function word. It is possible that sensitivity to transpositions, and therefore accuracy of response, may have been reduced in the latter relative to the former situation (see Wen et al., 2021).

2.4.3. Response time results

Overall, participants took just over 2 seconds to form their grammaticality decisions to the sentences (see Table 2.2). RTs showed a significant effect of transposition such that responses were longer for sentences with than without a transposition. The effect of final word grammaticality did not attain significance likely because of the very robust crossover interaction such that for sentences with a transposition, RTs were longer when they were grammatical than ungrammatical, while for sentences without a transposition, RTs were shorter when they were grammatical than ungrammatical. I, again, note that these RTs are somewhat longer than those reported in studies employing speeded GDTs. I also note that the pattern of effects for RTs is similar to that obtained by Mirault et al., (2018; 2020).

2.4.4. Eye movement measures results

The descriptive statistics for the eye movement measures are shown in Table 2.2. In particular, the means for the three first-pass reading measures (FFD, SFD and GD) for Pre-Target, Target Word 1, Target Word 2 and Post-Target are also visualised for convenience in Figures 2.2-2.5. GLMM summaries are shown in Table 2.3.

On the First Word the only measure for which significant effects were obtained, was RIn. Participants made more regressions to the First Word region when the two target words were not transposed compared with when they were transposed. There was also an effect of final word grammaticality such that readers made more regressions to the First Word region when the final word was grammatical compared with ungrammatical. Note, though, that both effects were qualified by an interaction between transposition and final word grammaticality.

The probability that readers regressed to fixate the First Word was increased for grammatically correct sentences, relative to sentences in all other conditions for which RIn measures were comparable. This result may initially appear somewhat surprising, in that, ordinarily, readers make more regressions and take longer to read sentences when they experience processing difficulty than when they do not. Presumably, a transposed word pair in a sentence, or an ungrammatical sentence final word (or both), should cause more processing difficulty than word pairs in their correct order or a grammatical sentence final word. However, recall that participants were required to decide whether the sentences were, or were not, grammatical. Under these circumstances, it is quite possible that the inflated RIn probabilities reflect re-reading in order to check that the sentence was actually grammatical, rather than being caused by disruption to reading resulting from a transposition or an ungrammatical sentence final word.

One final, important, point to make in relation to the results for the First Word region concerns the complete absence of any transposition effects on first-pass reading measures. That is to say, there was no evidence for any semantic or syntactic parafoveal-on-foveal effects associated with the transposed word pair manipulation. This result does not align with the predictions of the SWIFT model (Engbert, 2002; 2005) and its successor, SEAM (Rabe et al., 2024), both of which specify that the lexical characteristics of upcoming words in the parafovea should influence fixation durations. Although, as discussed previously (e.g., Section 1.6.2) under some circumstances, SWIFT and SEAM would predict that lexical processing may proceed in a serial-like manner (see Schad & Engbert, 2012). This is because upcoming word(s) in parafoveal vision receive little-to-no activation as attention is primarily focused on the currently fixated word to facilitate lexical identification. Therefore, the lack of early effects at the First-Word region does not necessarily contradict the SWIFT and SEAM frameworks.

For the Pre-Target, the pattern of results was quite similar to that obtained for the First Word region. There were no significant effects of word transposition on any first-pass reading times measures (see Table 2.2 and Figure 2.2). The only measures that showed robust effects were RIn and TVT. There was a significant effect of final word grammaticality on RIn and TVT, with increased regression rates into the Pre-Target region and longer reading times when the final word was grammatical than ungrammatical. There was also a significant interaction for RIn, such that the effect of transposition was very slightly larger when the sentence-final word was legal than illegal, and the probability that readers regressed to fixate the Pre-Target Word was greatest for grammatically correct sentences. These results further support the idea that readers re-read and checked legal sentences more than illegal sentences, and therefore, were able to make grammaticality decisions more rapidly when sentences contained a syntactic violation than when they did not. Furthermore, the results for the pre-target word again provide no evidence to support the parallel processing accounts of SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024) although, as discussed previously (see section 1.6.2; Schad & Engbert, 2012) the lack of any effects on these regions is not direct evidence against parallel processing.

Recall that the manipulations in this experiment did not involve changes to the first two words of the present sentences. That is to say, up until Target Word 1, the sentence was completely grammatical and made perfect sense in all conditions. However, Target Word 1 was the first word in the sentence that was (in some conditions) ungrammatical, and unsurprisingly, patterns of effects were markedly different at this word, where we obtained very early, substantive effects in the reading time measures. For Target Word 1, there was a significant effect of transposition on SFD, GD (see Figure 2.3), as well as GPT and TVT, with longer reading times when the two target words were transposed, compared to when they were not. There was also a main effect of transposition on RP, RIn, and ROut such that participants were more likely to refixate and regress into, and out of, Target Word 1 when the two targets were transposed compared with when they were not. All of these results indicate that the transposition of the two target words produced rapid and very substantive disruption to processing upon fixation of Target Word 1. Note also that this was the case in the context of participants correctly identifying that sentences containing a transposition were ungrammatical on over 90% of trials. Clearly, in the present study, participants were quite able to detect transpositions and to judge that sentences were ungrammatical when they fixated the first word of the transposed word pair. Furthermore, transpositions clearly had a rapid influence on eye movement behaviour. These results are entirely in line with serial processing accounts such as the E-Z Reader model (Reichle et al, 1998; Reichle, 2011) and Über Reader (Reichle, 2021; Veldre et al., 2020). These results might also be explained by parallel models such as SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024) in that a transposed word pair resulting in an ungrammatical sentence should produce disruption to reading when it is processed. This said, it is unclear why such effects did not occur earlier in the sentence if words are lexically identified whilst in the parafovea. Finally, these results are perhaps most problematic for the OB1 model (Snell, van Leipsig, et al., 2018) that specifies that words are identified in parallel and integrated together at a later stage of processing based on a spatiotopic map. If this was the case, then it is unclear why participants detected word transpositions on between 92-99% of trials, nor is it clear why transpositions caused substantive and immediate disruption to eye movements upon fixation of the first word of the transposed word pair.

With respect to final word grammaticality, there was no effect on any of the first-pass measures on Target Word 1 (see Table 2.2 and Figure 2.3). That is to say, there was no evidence to suggest that an ungrammatical final word had a disruptive influence on initial processing at Target Word 1. This, similar to the lack of transposition effects for early measures on First Word and Pre-Target does not provide support for the parallel assumptions of SWIFT and SEAM in particular. However, as discussed above, it is not necessarily inconsistent with these models given that there is an absence of an effect (e.g., see Schad & Engbert, 2012). Note, that for later measures (e.g., RIn and TVT), there were robust effects of final word grammaticality. Participants were more likely to regress into Target Word 1, and spend longer re-reading Target Word 1, when the final word was grammatical than ungrammatical. This pattern of late effects is very comparable to the effects obtained for the First Word and Pre-Target regions and presumably reflects increased checking for grammatical sentences as suggested earlier. Beyond these results, there were no significant interactions for any measure for Target Word 1.

On Target Word 2 there was a significant effect of transposition on GPT, and TVT, with longer reading times when the two targets were transposed, compared to not transposed. This pattern of effects is very similar to that observed for Target Word 1 and suggests disruption to eye movements due to the word transposition and resulting ungrammaticality. There was also an effect of transposition on RIn and ROut. Participants were less likely to regress into, and more likely to regress out of Target Word 2 when the two targets were transposed compared with when they were not transposed.

Additionally, on Target Word 2, there was a significant effect of final word grammaticality for GPT and ROut, with longer GPT and a higher probability of regressions from the Target Word 2 region when the final word was grammatical compared to when it was ungrammatical. It is important to note that these findings indicate that the grammaticality of the final word was detected from the parafovea and influenced processing on the Target Word 2 region. The presence of this effect is only consistent with SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024). This is because both models would predict that the final word grammaticality should influence processing on Target Word 2 given that words are identified and integrated in parallel. A further important point in relation to these effects is that they are in the opposite direction to what I might have expected. That is to say, readers made fewer regressions out of the region when the sentence-final word was ungrammatical than grammatical, though note that these effects were again qualified by an interaction. I also observed an effect of the final word grammaticality on RIn for the Target Word 2 region, with readers making more regressions into this region when the final word was ungrammatical than when it was grammatical. Again, though, this effect was qualified by a significant interaction between final word grammaticality and transposition.

Let us turn to the interactive patterns of effects. Broadly, there were three notable aspects of the results. First, for the GD (see Table 2.2 and Figure 2.4) there was a numerical pattern that matched the robust pattern of effects that were observed for GPT, TVT and ROut measures. I found that reading times were longest and regressions to earlier regions of the sentence were increased for sentences with a transposition and a legal final word relative to all the other conditions. This pattern of effects, presumably, simply reflects readers experiencing disruption when processing the second word of the transposed word pair (in quite a similar way to how they experienced disruption at the Target Word 1 region). The second noteworthy aspect of the interactive effects relates to the ROut and the GPT measures specifically. Readers were less likely to regress from the Target Word 2 region to inspect an earlier region of the sentence, and if they did make a regression, they also spent less time re-reading earlier portions of the sentence when those sentences did not contain a transposition but did contain an ungrammatical sentence-final word. This finding suggests a parafoveal sensitivity to the grammaticality of the word that appeared at the end of the sentence, and in the absence of a transposition, if the word was ungrammatical, readers were less likely to make a leftward saccade (and, therefore, more likely to make a rightward saccade – see the fixation probability effects for the final word of the sentence for a pattern of effects that complements this suggestion). That is, it appears that when the final word was ungrammatical, and readers detected this ungrammaticality whilst fixating the preceding word, their point of fixation was drawn to the ungrammaticality presumably to allow the reader to check that the word was indeed ungrammatical. What is striking is that on detecting a parafoveal ungrammaticality, readers did not exhibit immediate disruption to processing at the point that the ungrammaticality was detected (e.g., increased fixation durations on the Target Word 2 region, increased probability of regressions to another word to the left and increased re-reading of early portions of the sentence). Instead, they moved their eyes forward in the text to verify the ungrammaticality. The third notable interactive pattern may be related to this. For sentences without a transposition but with an ungrammatical final word, I also found that readers were substantially more likely to make a regression from the final word back to Target Word 2 (i.e., the RIn measure). The RIn pattern of effects fits with the suggestion that after readers had verified that the final word was ungrammatical by fixating it, they were then more likely to regress from the final word to reinspect the Target Word 2 region. To summarise this set of results, it appears that readers experienced disruption at Target Word 2 when sentences contained a transposition. However, when sentences did not contain a transposition, but did have an ungrammatical final word, readers detected the ungrammaticality parafoveally. Despite this, detection of the ungrammaticality did not cause readers to experience difficulty at Target Word 2, but instead caused them to move their eyes forward in order to fixate the final word of the sentence. Furthermore, when this word was indeed ungrammatical, readers were then more likely to make a regression back into the Target Word 2 region. I will consider these aspects of the results further in the Discussion.

Let us next consider the results for the Post-Target word (see Figure 2.5). Here there was a significant effect of transposition for SFD, GD, GPT, TVT, with longer reading times when the two targets were not transposed versus transposed; and on FP, RP, and ROut, with a lower probability to fixate, refixate and to regress out of the Post-Target when the two targets were transposed versus not transposed. I also found an effect of final word grammaticality on SFD, GD, and TVT, with longer reading times when the final word was ungrammatical than grammatical, as well as on FP and RP with a higher probability to fixate and refixate the sentence final word when it was ungrammatical than grammatical. The effects of word transposition and final word grammaticality were qualified by interactions that were robust for SFD, GD, RP, GPT, TVT and patterned numerically for FFD, FP and ROut. In the reading time measures, these interactions were largely driven by longer times for the Post-Target word when it was ungrammatical, and the sentences did not contain a transposition. Note that this was particularly the case for GPT. Readers spent longest re-reading sentences with an ungrammatical final word but no transposition (and note also that this effect complements the RIn effect observed for the Target Word 2 region). These results are consistent with participants experiencing disruption to processing when they encountered an ungrammatical word for the first time. In the three other conditions, readers had either already encountered an ungrammaticality (i.e., a transposed word pair) earlier in the sentence, and if so had already detected the anomaly and reacted to it (by taking longer to read the words, regressing and re-reading early portions of the sentence). Alternatively, they had processed a sentence that was perfectly grammatical up to the final word and had, therefore, not experienced difficulty until this point in the sentence. Two other differences seem to contribute significantly to the interactive effects, namely, the relatively low rate of regressions from the final word and the reduced GPTs when the sentence contained both a transposed word pair and an ungrammatical final word relative to all other conditions. As suggested previously, it seems very likely that values for these measures were reduced because participants had already experienced disruption to processing upon detecting the ungrammaticality created by the transposed word pair earlier in the sentence. To be clear, at the Post-Target word, readers had almost certainly already detected the ungrammaticality associated with the transposition, and consequently, when the Post-Target word provided an additional indication that the sentence was ungrammatical, this had a far reduced disruptive influence on processing (in that it provided a redundant cue to an ungrammatical sentence).

In fact, given that overall grammaticality decision times were shortest for the grammatically legal sentences and sentences containing a transposition as well as a grammatically illegal final word, it might even be the case that having two cues to ungrammaticality actually reduced disruption by facilitating readers in forming their decision.

Table 2.2.

Descriptive statistics for all measures

Measure/ Condition	Grammatical and Not Transposed	Grammatical and Transposed	Ungrammatical and Not Transposed	Ungrammatical and Transposed
Accuracy	93 (26)	92 (27)	91 (28)	99 (11)
RTs	2148 (781)	2272 (830)	2246 (797)	2160 (784)
	First Word			
	Not Transposed		Transposed	
FFD	182 (54)		180 (52)	
SFD	179 (53)		177 (52)	
GD	203 (94)		206 (101)	
RP	12 (32)		14 (34)	
	Grammatical and Not Transposed	Grammatical and Transposed	Ungrammatical and Not Transposed	Ungrammatical and Transposed
TVT	242 (129)	232 (129)	224 (130)	225 (123)
RIn	25 (43)	15 (36)	16 (36)	16 (36)
	Pre-Target			
	Not Transposed		Transposed	
FFD	197 (53)		194 (53)	
SFD	204 (53)		201 (53)	
GD	240 (88)		238 (89)	
SP	7 (26)		8 (27)	
RP	24 (43)		23 (42)	
ROut	4 (20)		5 (21)	
GPT	251 (102)		249 (105)	
	Grammatical and Not Transposed	Grammatical and Transposed	Ungrammatical and Not Transposed	Ungrammatical and Transposed
TVT	355 (187)	348 (201)	320 (187)	322 (194)
RIn	42 (49)	36 (48)	25 (43)	28 (45)
	Target Word 1			
FFD	219 (63)	223 (69)	217 (60)	222 (69)
SFD	230 (60)	245 (79)	229 (65)	241 (76)
GD	255 (90)	276 (112)	257 (93)	269 (106)
SP	13 (34)	13 (33)	13 (34)	12 (33)
RP	17 (38)	21 (41)	18 (39)	20 (40)
ROut	6 (24)	8 (27)	5 (21)	8 (28)
GPT	274 (117)	305 (147)	271 (119)	296 (145)
TVT	397 (231)	463 (256)	371 (220)	434 (247)
RIn	44 (50)	51 (50)	36 (48)	45 (50)
	Target Word 2			
FFD	249 (82)	253 (85)	251 (83)	251 (91)
SFD	282 (95)	300 (99)	287 (96)	295 (101)
GD	328 (145)	346 (171)	331 (148)	329 (156)
SP	8 (27)	7 (25)	9 (29)	7 (25)
RP	29 (45)	31 (46)	29 (45)	28 (45)
ROut	24 (43)	34 (47)	13 (34)	29 (45)
GPT	430 (273)	537 (317)	391 (221)	470 (265)
TVT	467 (237)	542 (268)	516 (255)	502 (254)
RIn	30 (46)	27 (44)	46 (50)	29 (46)
	Post-Target			
FFD	244 (95)	243 (106)	257 (95)	237 (95)
SFD	251 (103)	246 (113)	273 (105)	250 (111)
GD	309 (160)	305 (176)	368 (190)	304 (165)
FP	79 (41)	72 (45)	90 (30)	84 (37)
RP	26 (44)	22 (42)	38 (48)	23 (42)
ROut	69 (46)	63 (48)	70 (46)	57 (50)
GPT	764 (576)	786 (639)	844 (554)	676 (554)
TVT	380 (244)	367 (240)	456 (242)	369 (218)

Note. Values for FFD, SFD, GD, GPT, and TVT are given in milliseconds while values for Accuracy, SP, RP, FP, RIn, and ROut are given in percentages.

Table 2.3.

Fixed effects estimates from the generalised mixed effects models.

Measure/ Condition	Intercept				Transposition (Transposed-Not Transposed)				Final Word Grammaticality (Grammatical-Ungrammatical)				Interaction			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
Accuracy	3.89	0.22	**18.07**	[3.47, 4.31]	1.24	0.39	**3.16**	[0.47, 2.01]	-1.55	0.45	**-3.41**	[-2.44, -0.66]	-2.67	0.82	**-3.25**	[-4.28, -1.06]
RTs	2408.47	6.45	**373.69**	[2395.83,2421.10]	37.71	8.37	**4.50**	[21.30,54.13]	18.55	8.27	2.24	[2.34,34.75]	233.83	6.16	**38.92**	[227.75,251.91]
	First Word															
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z) -value*	*Conf Int*	*β*	*SE*	*t(z) -value*	*Conf Int*	*β*	*SE*	*t(z) -value*	*Conf Int*
FFD	182.31	6.40	**28.48**	[167.98,196.63]	-2.44	1.88	-1.30	[-6.19,1.31]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	192.69	8.61	**22.38**	[175.81,209.56]	-2.69	2.27	-1.18	[-7.14,1.76]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	207.02	9.21	**22.48**	[188.96,225.07]	2.23	2.84	0.79	[-3.33,7.79]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
RP	-3.38	0.28	**-12.25**	[-3.92, -2.84]	0.03	0.13	0.25	[-0.23,0.29]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	234.49	7.04	**33.29**	[220.68,248.29]	-6.58	4.59	-1.43	[-15.57,2.41]	10.32	4.77	2.17	[0.98,19.66]	-10.26	7.21	-1.42	[-24.38,3.86]
RIn	-1.85	0.19	**-9.81**	[-2.22, -1.48]	-0.31	0.11	**-2.96**	[-0.52, -0.11]	0.49	0.13	**3.67**	[0.23,0.75]	-0.66	0.21	**-3.15**	[-1.07, -0.25]
	Pre-Target															
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	200.37	5.10	**39.32**	[190.39,210.36]	-3.14	2.46	-1.27	[-7.97,1.69]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	210.99	6.02	**35.04**	[199.19,222.79]	-2.14	3.51	-0.61	[-9.01,4.74]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	242.86	5.89	**41.24**	[231.32,254.40]	-2.34	3.78	-0.62	[-9.74,5.06]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SP	-3.39	0.26	**-13.03**	[-3.90, -2.88]	-0.17	0.2	-0.87	[-0.55,0.21]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
RP	-1.49	0.17	**-8.65**	[-1.82, -1.15]	-0.01	0.08	-0.17	[-0.18,0.15]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ROut	-3.60	0.19	**-18.78**	[-3.98, -3.22]	0.10	0.17	0.59	[-0.23,0.42]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GPT	257.44	6.92	**37.22**	[243.89,271.00]	-1.37	4.2	-0.33	[-9.60,6.85]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	350.19	9.54	**36.72**	[331.50,368.88]	2.06	9.96	0.21	[-17.46,21.59]	34.85	7.02	**4.97**	[21.09,48.61]	0.57	9.15	0.06	[-17.36,18.50]
RIn	-0.96	0.20	**-4.84**	[-1.35, -0.57]	0.01	0.11	0.07	[-0.22,0.23]	0.79	0.10	**7.62**	[0.59,1.00]	-0.53	0.16	**-3.22**	[-0.85, -0.21]
	Target Word 1															
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	223.49	5.49	**40.72**	[212.74,234.25]	4.86	3.88	1.25	[-2.75,12.46]	0.88	2.75	0.32	[-4.51,6.26]	-3.02	5.29	-0.57	[-13.38,7.34]
SFD	241.09	6.98	**34.54**	[227.41,254.77]	16.26	6.23	**2.61**	[4.04,28.48]	5.53	5.32	1.04	[-4.90,15.96]	3.76	10.43	0.36	[-16.68,24.19]
GD	265.69	5.37	**49.53**	[255.18,276.21]	18.28	5.19	**3.52**	[8.11,28.46]	1.21	4.28	0.28	[-7.18,9.60]	5.51	6.32	0.87	[-6.87,17.90]
SP	-2.60	0.23	**-11.40**	[-3.05, -2.15]	-0.06	0.11	-0.57	[-0.27,0.15]	0.06	0.10	0.59	[-0.14,0.26]	0.12	0.21	0.58	[-0.29,0.52]
RP	-1.71	0.14	**-11.84**	[-1.99, -1.43]	0.26	0.09	**2.91**	[0.09,0.44]	0.02	0.09	0.20	[-0.16,0.19]	0.10	0.18	0.58	[-0.25,0.45]
ROut	-2.99	0.15	**-19.66**	[-3.29, -2.70]	0.54	0.14	**3.96**	[0.27,0.80]	0.14	0.14	1.02	[-0.13,0.40]	-0.35	0.27	-1.31	[-0.88,0.18]
GPT	290.63	5.92	**49.13**	[279.03,302.22]	32.14	5.95	**5.40**	[20.47,43.80]	5.75	6.67	0.86	[-7.32,18.83]	7.64	9.54	0.80	[-11.05,26.33]
TVT	429.27	5.65	**75.96**	[418.19,440.35]	74.59	6.03	**12.38**	[62.78,86.40]	29.34	5.98	**4.91**	[17.62,41.05]	8.00	9.69	0.83	[-10.99,26.98]
RIn	-0.32	0.15	-2.17	[-0.61, -0.03]	0.41	0.09	**4.78**	[0.24,0.58]	0.38	0.08	**4.56**	[0.22,0.54]	-0.20	0.15	-1.34	[-0.49,0.09]
	Target Word 2															
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	255.51	5.67	**45.08**	[244.40,266.62]	2.64	4.41	0.60	[-6.00,11.28]	-0.65	4.47	-0.15	[-9.41,8.10]	-0.72	6.72	-0.11	[-13.90,12.45]
SFD	294.07	8.59	**34.23**	[277.23,310.90]	14.19	9.29	1.53	[-4.02,32.39]	-2.13	9.37	-0.23	[-20.49,16.24]	3.84	13.80	0.28	[-23.21,30.89]
GD	338.09	6.95	**48.66**	[324.47,351.71]	11.16	6.31	1.77	[-1.20,23.52]	7.71	6.38	1.21	[-4.79,20.20]	15.58	6.62	2.35	[2.60,28.56]
SP	-3.04	0.20	**-15.34**	[-3.43, -2.66]	-0.27	0.18	-1.45	[-0.62,0.09]	-0.05	0.17	-0.26	[-39,0.30]	0.02	0.36	0.05	[-0.69,0.73]
RP	-0.98	0.11	**-9.29**	[-1.19, -0.78]	0.05	0.07	0.65	[-0.10,0.19]	0.09	0.08	1.10	[-0.07,0.24]	0.17	0.14	1.18	[-0.11,0.45]
ROut	-1.33	0.13	**-10.11**	[-1.59, -1.07]	0.85	0.11	**7.99**	[0.64,1.06]	0.59	0.08	**6.98**	[0.42,0.75]	-0.59	0.17	**-3.53**	[-0.93, -0.26]
GPT	466.95	6.75	**69.19**	[453.73,480.18]	100.98	8.55	**11.81**	[84.22,117.74]	63.35	6.36	**9.96**	[50.88,75.81]	34.58	7.08	**4.88**	[20.69,48.46]

TVT	520.2	7.76	**67.02**	[504.99,535.41]	33.36	5.44	**6.13**	[22.70,44.02]	-6.94	7.30	-0.95	[-21.24, 7.36]	93.21	9.69	**9.62**	[74.22,112.20]
RIn	-0.89	0.14	**-6.29**	[-1.16, -0.61]	-0.5	0.08	**-6.57**	[-0.65, -0.35]	-0.49	0.08	**-5.71**	[-0.63, -0.31]	0.63	0.15	**4.16**	[0.33,0.93]
	Post-Target															
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	244.96	7.61	**32.17**	[230.04,259.88]	-13.36	6.03	-2.22	[-25.17, -1.55]	-6.58	5.80	-1.13	[-17.95,4.79]	20.27	8.42	2.41	[3.77,36.77]
SFD	269.21	8.21	**32.81**	[253.13,285.29]	-21.72	6.65	**-3.27**	[-34.75, -8.69]	-21.00	7.71	**-2.72**	[-36.12, -5.89]	24.19	8.89	**2.72**	[6.77,41.61]
GD	318.46	7.76	**41.03**	[303.24,333.67]	-41.41	7.69	**-5.39**	[-56,48, -26.34]	-38.49	6.07	**-6.34**	[--50.39, -26.59]	65.32	7.36	**8.88**	[50.91,79.74]
FP	2.03	0.20	**10.01**	[1.63,2.43]	-0.48	0.13	**-3.70**	[-0.73, -0.22]	-0.97	0.13	**-7.41**	[-1.23, -0.72]	0.30	0.20	1.52	[-0.09,0.68]
RP	-1.32	0.15	**-8.60**	[-1.62, -1.02]	-0.59	0.11	**-5.41**	[-0.80, -0.37]	-0.46	0.12	**-3.72**	[-0.70, -0.22]	0.64	0.22	**2.87**	[0.20,1.08]
ROut	0.72	0.17	**4.31**	[0.39,1.05]	-0.52	0.08	**-6.25**	[-0.69, -0.36]	0.11	0.11	0.96	[-0.11,0.33]	0.40	0.17	2.42	[0.08,0.73]
GPT	747.06	7.42	**100.67**	[732.52,761.61]	-85.21	8.80	**-9.68**	[-102.46, -67.96]	-6.3	8.34	-0.76	[-22.65,10.04]	175.39	7.12	**24.63**	[161.44,189.35]
TVT	386.62	7.79	**49.61**	[371.35,401.89]	-56.45	5.94	**-9.50**	[-68.10, -44.80]	-55.93	6.52	**-8.57**	[-68.72, -43.15]	86.85	9.95	**8.73**	[67.34,106.36]

Note. Significant terms are presented in bold, and terms approaching significance are underlined. All *p* values were adjusted via the Bonferroni correction with the summary(glht(<modelName>), test = adjusted(“bonferoni”)) function. N/A signifies that the factor or interaction was not considered in the model.

^Confidence intervals (Conf Int) are calculated with the *confint* function in R with method = *Wald* at the 2.5 percentile and 97.5 percentile

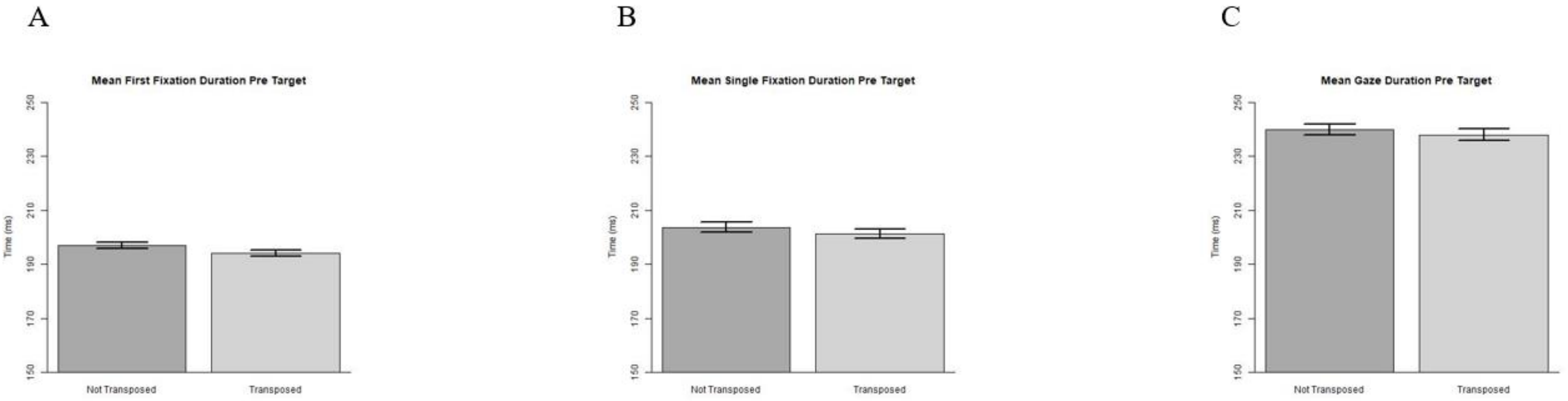


Figure 2.2. Mean first-pass reading times on Pre-Target

Panel A shows the mean first fixation duration on the Pre-Target region; Panel B shows the mean single fixation duration on the Pre-Target region; Panel C shows the mean gaze duration on the Pre-Target region.

Error bars represent standard errors.

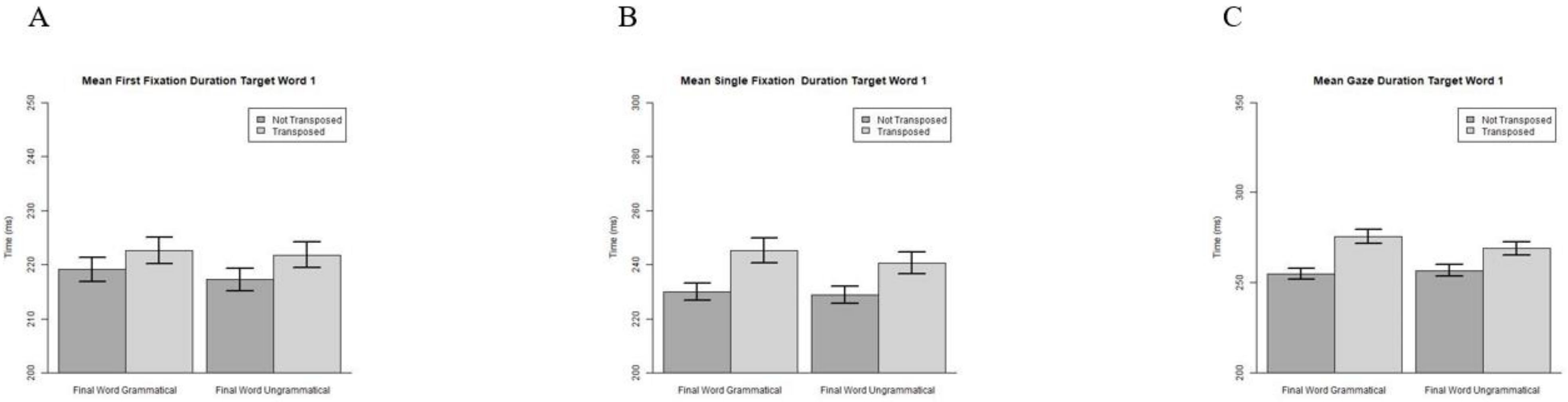


Figure 2.3. Mean first-pass reading times on Target Word 1

Panel A shows the mean first fixation duration on the Target Word 1 region; Panel B shows the mean single fixation duration on the Target Word 1 region; Panel C shows the mean gaze duration on the Target Word 1 region.

Error bars represent standard errors.

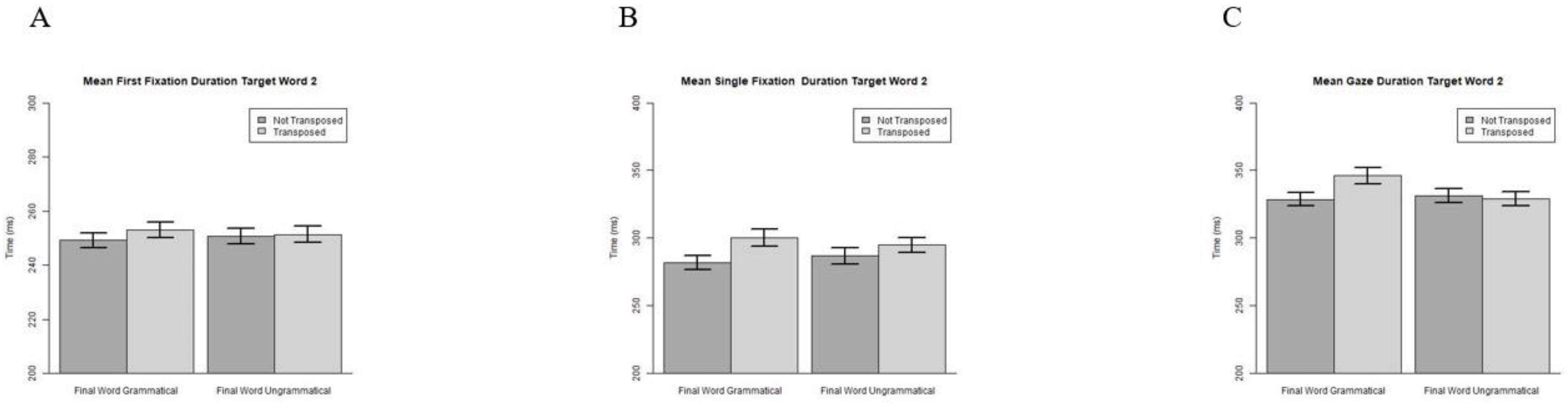


Figure 2.4. Mean first-pass reading times on Target Word 2

Panel A shows the mean first fixation duration on the Target Word 2 region; Panel B shows the mean single fixation duration on the Target Word 2 region; Panel C shows the mean gaze duration on the Target Word 2 region.

Error bars represent standard errors.

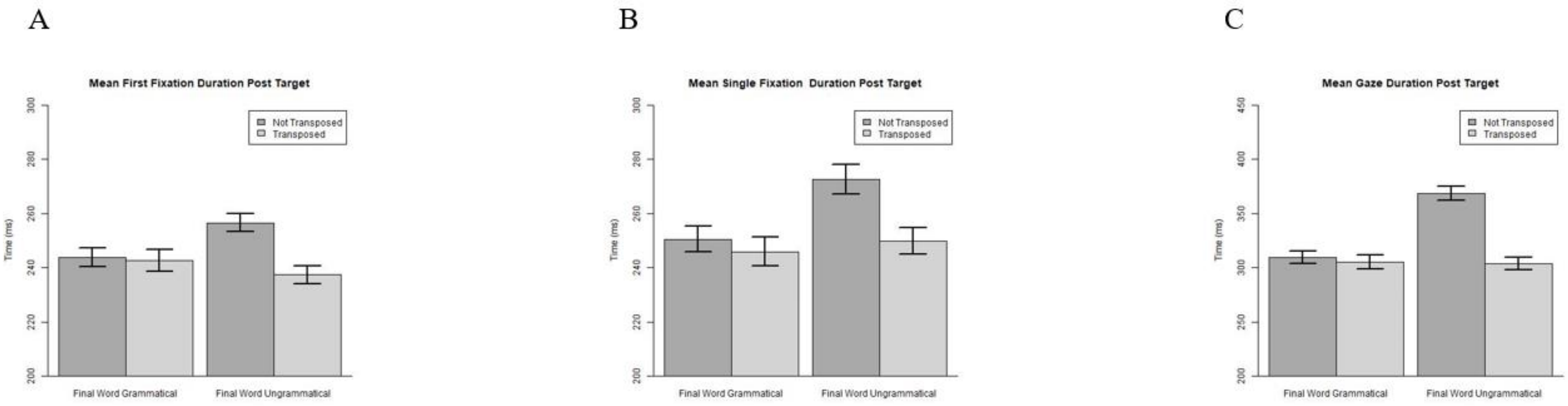


Figure 2.5. Mean first-pass reading times on Post-Target

Panel A shows the mean first fixation duration on the Post-Target region; Panel B shows the mean single fixation duration on the Post-Target region; Panel C shows the mean gaze duration on the Post-Target region.

Error bars represent standard errors.

2.5. Discussion

In the present experiment, I investigated how transposed word pairs and sentence-final words that were ungrammatical affected eye movement behaviour as participants processed sentences in order to make a grammaticality decision. My motivation for the experiment was to examine these two influences through their orthogonal manipulation and to use a set of stimuli that had been well-normed to ensure that both manipulations resulted in sentences that participants perceived to be ungrammatical. Particularly, I sought to understand how frequently, and at what point during processing, participants first detected ungrammaticalities and whether, and how, both cues to ungrammaticality affected eye movements as the sentences were read. In doing this, I also evaluated which of the current models of eye movement behaviour best accounted for the present findings.

In some respects, the results from the present study were very clear and straightforward, whilst in other respects, it is probably fair to say that the patterns of results were complex. Before I consider each of the present findings in turn, it is perhaps worth commenting on some general characteristics of the present experiment that likely contributed to complexities in the findings. The current experiment was based on a series of original experiments that first showed the *Transposed-Word* effect (Mirault et al., 2018; 2020). These original experiments used very similar linguistic stimuli that took the form of 5-word sentences. The third and fourth words of the sentence were potentially transposed, while the final, fifth word of the sentence could either be grammatical or ungrammatical but only when preceded by a transposed word pair. Because I wished to use stimuli that were directly comparable to the types of stimuli that were used in the original experiments, I adopted sentences of this form here. However, it is important to note that there are a number of reasons why stimuli with this type of construction might not be the most ideal to use in an eye movement experiment that investigates the time course of ungrammaticality detection. First, the sentences are short, which means that readers will not make too many eye movements as they read those sentences, and of course, this is not ideal because eye movement behaviour is the dependent measure in this experiment. More important, however, is the fact that the two manipulations of ungrammaticality in the sentence, namely, the word pair transposition and the grammaticality of the sentence-final word take place immediately adjacent to each other in the sentence. Given that sentential ungrammaticality produces significant and extended disruption to eye movement behaviour during reading, this means that effects associated with the first ungrammaticality manipulation (the transposed word pair) will almost certainly spillover and affect fixations made on the following word, that is, the sentence-final word which itself was manipulated for grammaticality. Often in eye movement experiments that seek to investigate the influence of two different variables on processing, experimenters design their stimuli to ensure that several words appear in a sentence or a text between the portions of the sentence that will likely produce disruption to processing. This is done to make it possible to tease apart and separately observe particular effects in the eye movement record that are associated specifically with each variable. Thus, in the present experiment, the fact that the sentences were only 5 words long in total and that the regions of text associated with the two manipulations of sentence ungrammaticality themselves comprised three of the five words in the sentence, and that these two regions were adjacent to each other, meant that there was a significant degree of congestion with respect to the influences of the variables that were manipulated. Furthermore, ordinarily in eye movement studies investigating reading, experimenters most often choose not to implement critical experimental manipulations on sentence final words since for such words readers do not have an opportunity to move their eyes to a word to the right and thereby face a decision of regressing or remaining fixated on the final word in the sentence. Furthermore, fixations on the final word of a sentence are also associated with sentence wrap-up effects (e.g., Warren et al., 2009) and such effects might likely cloud any experimental effects associated with manipulations of the final word. It is very likely, therefore, that the effects obtained in the present experiment may not have been as clear as might have been preferred had the sentences been longer and the two manipulations had been spaced apart to a greater degree within the sentence (e.g., 12-word sentences with the transposed word pair appearing at positions 4 and 5 and the ungrammatical word appearing at position 10 within the sentence). Stimuli with such structure may have produced less congested effects and would have allowed for more effective examination of the time course of the two sources of influence on processing. And to reiterate, the reason that I did not adopt sentences that were more optimally formed in these regards was because I wished to use stimuli that allowed direct comparison with previous experiments.

Regardless of the particular form of the sentences used here, it was certainly the case that the experimental manipulations produced very robust effects, and this was at least in part due to the pre-screening procedures that were adopted prior to conducting the eye movement experiment. Recall, unlike almost all previous studies investigating *Transposed-Word* effects in reading, I assessed the degree to which subjects judged the present sentences as natural when the target word pairs were, or were not, transposed and when the sentence final word was, or was not, grammatically illegal. And in line with my expectations, readers rated sentences as unnatural when they contained a transposition, or an ungrammatical final word, and particularly when the sentences contained both. In this way, I could be sure that the transposed word pair and sentence final word manipulations were such that they caused readers to assess the stimuli as ungrammatical as required.

Furthermore, as mentioned earlier, it was the case that aspects of the present results were very clear. To briefly summarize, participants had high levels of accuracy in this experiment, making correct grammaticality decisions on over 90% of occasions in all conditions, and performing almost perfectly when sentences contained two cues indicating that they were ungrammatical. These grammaticality decision accuracy rates are substantially higher than those that have been previously reported; for example, for the sentences containing only a word transposition, between 84% (in a laboratory experiment) and 88% (in an on-line experiment) in Mirault et al. (2018), 84% in Mirault et al. (2020), and between 76% (when the transposed word pair contained two short open-class words) and 91% (when the transposed word pair contained one open-class and one closed-class word and one word was short while the other was long) in Huang and Staub (2021a). The increased accuracy rates in the current study very likely reflect the fact that the stimuli in the present experiment were screened to ensure that the word transposition and sentence-final word manipulations both provided an unambiguous cue to sentence ungrammaticality. Whether these manipulations were quite so effective in (at least some of the) previous studies is open to question. A second reason for the reduced error rates in the present study is that the present experimental task required participants to read the sentences carefully and to understand them to the best of their ability. Previous studies have often required participants to make speeded grammaticality decisions which may have contributed to participants making more errors (see Liu et al., 2021). That said, similarly to Liu and colleagues (2021) there was still a *Transposed-Word* effect on accuracy such that participants were better at correctly judging sentences with two syntactic violations as ungrammatical (M = 99%) versus sentences containing a word transposition only (M = 92%). These findings further support the notion that *Transposed-Word* effects can be observed in an unspeeded grammaticality judgment task and are consistent with existing literature.

Considering the overall RTs results, these were found to be shorter when sentences were grammatical than when sentences were ungrammatical. However, when participants decided that sentences were ungrammatical, their decisions were slower when sentences contained one cue to their ungrammaticality (either a transposed word pair, or an ungrammatical final word) than when they contained both cues (both a transposed word pair and an ungrammatical final word). This is consistent with Mirault et al. (2018; 2020), as they also found RTs to be shorter for sentences containing both a transposition and a final ungrammatical word than sentences containing only the word transposition. Note, though, that in the current study RTs were somewhat longer (approximately 2.2-2.3s) than has been reported in previous studies (approximately 1.3-1.8s, Mirault et al., 2018; 1.7s, Mirault et al., 2020). The patterns observed for RTs fit with those observed for accuracy and the *Transposed-Word* effects previously reported in the literature (e.g., Mirault et al., 2018, 2020).

Next, I will consider the eye movement and reading time data for each region of the sentences. When participants started reading the sentences, they did so without any difficulties over the first two words and this was the case for sentences in all conditions. However, as soon as participants fixated the first word of a transposed word pair (but not before this point), they very rapidly experienced disruption to processing relative to when the word pair was not transposed. Thus, at the third word of the 5-word sentence (Target Word 1 region), there was disruption to first pass reading caused by the third and fourth word having their positions switched. Note also that there was no influence of the grammaticality of the final word of the sentence at this point. These aspects of the results are straightforward and very clear. From the grammaticality pre-screen data, it is clear that when the stimuli contained a transposed word pair, they provided an unambiguous cue to their ungrammaticality at the Target Word 1 region. This explains why participants detected the ungrammaticality on almost all of the trials, and the first point at which they did this was when they fixated the third word in the sentence, that is, the first of the transposed word pair (the Target Word 1 region). In this regard, these results fit perfectly with a serial processing account in that there was no influence of the transposed word pair prior to the fixation of the first word of the pair, nor was there a parafoveal influence of the grammaticality of the final word in the sentence at the Target Word 1 region. It is not immediately obvious that the results from the first three regions of the sentence are in accord with the OB1 Reader model. The OB1 Reader model specifies that words are identified in parallel and that any ungrammaticality arising due to words appearing out of order should not produce immediate disruption to reading. Recall that OB1 Reader stipulates that words are encoded in relation to a spatiotopic map and through this mechanism, readers are able to process words when they do not appear in their correct grammatical order. Taking grammaticality decisions and reading time effects together, there should be two consequences of this specification that did not occur in RTs, nor in relation to the first pass eye movement measures for the first three regions of the sentence. First, readers should have failed to detect ungrammaticalities arising due to words appearing out of order, and second, readers should not have exhibited disruption to processing when they encountered transposed words. The current results show that readers almost always detected ungrammaticalities due to word transpositions with accuracy rates of over 90%. Furthermore, readers showed rapid disruption to processing when they fixated the first word of a transposed word pair. To my mind, therefore, the first-pass reading time results for the Target Word 1 of the sentences are inconsistent with the account offered by the OB1 Reader model. Finally, the SWIFT and SEAM models suggest that words are identified in parallel during reading and therefore parafoveal-on-foveal effects might be expected. That is to say, readers might identify ungrammatical words when those words are in the parafovea, that is, before they are fixated. This clearly did not happen. Thus, the first-pass reading times for the early portions of the sentence are inconsistent with the account offered by SWIFT and SEAM.

It is notable that the first-pass reading time results at the point of ungrammaticality are quite consistent with those reported by Huang and Staub (2021a). Although Huang and Staub reported lower accuracy rates for their sentences containing word pair transpositions than those obtained in the present study, they found that when readers did identify that sentences were ungrammatical, there was rapid disruption to processing in the eye movement record, in accord with the current results. It is likely that grammatical decision accuracy rates in their study were lower, specifically for sentences containing two short words that were often function words (2-4 letters long), because participants (at least in Experiment 1) were required to make speeded judgments.

Let us next consider the results from the fourth word of the sentence (Target Word 2 region). Here the results became somewhat more complicated. The disruptive effects of the transposition that were observed at the Target Word 1 region persisted through the Target Word 2 region – this aspect of the current findings is straightforward to explain. Both words comprising the transposed word pair provided a cue to ungrammaticality, so it is unsurprising that similar effects of transposition were observed here as occurred at the first word of the pair, one word earlier. There were also effects of the grammaticality of the sentence final word at this point, but more importantly, there were interactive effects of transposition and grammaticality of the final word that indicated readers were processing the final word to a significant degree prior to fixating it. That is to say, I observed effects showing a parafoveal-on-foveal influence of final word grammaticality. Specifically, when sentences did not contain a transposition, readers were less likely to regress and re-read the sentence when the final word was ungrammatical compared with when it was grammatical. In other words, when the final word of the sentence provided the only cue to its ungrammaticality, readers were sensitive to this and made the decision to fixate that word rather than making a regression to re-read the sentence (the fixation probability results for the final word also reflected this pattern). To be absolutely clear, it appears that readers were sensitive to the grammatical status of the final word before fixating it, and when it was ungrammatical, the ungrammaticality did not cause disruption to reading, but instead, it caused readers' eyes to be drawn to fixate it. This result was complemented by the finding that in this experimental condition, after fixating the final word in the sentence, readers were more likely to make a regression to earlier regions of the sentence to re-read them. It is these aspects of the results for the Target Word 2 region that I consider to be somewhat more complex to explain. According to a serial processing account such as E-Z Reader and Über Reader, there should be no parafoveal sensitivity to the grammatical status of the sentence final word prior to its fixation. The interactive effects at the Target Word 2 region are therefore inconsistent with a serial account. In contrast, according to a parallel account such as the SWIFT and SEAM models, readers should be sensitive to lexical properties of words prior to their fixation. However, it is not clear why the ungrammaticality of the final word would not have caused disruption to processing, but instead caused readers to be more likely to make a saccade to fixate the sentence final word. Finally, and similarly, whilst the OB1 Reader account of processing specifies that words are identified in parallel, the model also quite clearly specifies that such parallel processing should not result in discernible parafoveal-on-foveal effects (see Zang et al., 2023 for relevant discussion). Thus, this particular finding does not seem to fit neatly with any of the accounts put forward by existing computational models.

I must be transparent that this aspect of the present findings was certainly not something that I predicted, and nor was the explanation for it immediately self-evident to me. This result also caused additional confusion regarding why there were no such comparable effects when readers fixated the second word in the sentence (i.e., the Pre-Target word region) and the sentence contained a transposed word pair in the parafovea. In such a situation, readers would also have a word to the right of the fixated word that was ungrammatical. Despite the similarity in these circumstances, there was no evidence of any parafoveal ungrammaticality detection. The absence of similar effects earlier in the sentence led me to the view that the final word effects could have arisen particularly because the ungrammatical word was both sentence and line final. Recall that in eye movement experiments, researchers most often avoid placing critical regions at the end of sentences or in line final positions to ensure that readers have an opportunity to move their eyes rightward (in alphabetic languages like English) to fixate upcoming words and to avoid the possibility of experimental effects being clouded by sentence wrap-up and return sweep effects. Whilst I have no ready explanation for the final word effects that were observed, I suspect that the way in which readers process line final words, particularly when they are likely to regress to re-read earlier portions of a sentence, or make a return sweep, or make a large leftward saccade in order to prepare for the next trial of the experiment, may be different to the way in which the other words of the sentence are processed. For example, it may be the case that readers parafoveally process line final words to gain a sense of whether the upcoming word appears to be visually consistent with predictable candidates. If the upcoming word does look somewhat predictable, then it might likely be the case that a leftward eye movement is launched without a direct fixation to the line final word. In contrast, when the line final word does not resemble a predictable word (i.e., it is a word that is unlikely to appear in the sentence based on context, e.g., an ungrammatical word), then readers may be more likely to make a saccade to it in order to fixate it prior to making a subsequent leftward saccade. If this was the case, then it is likely that readers quite frequently fail to fixate line final words, and they do this to a greater degree than they skip line internal words. Clearly, further research is required to evaluate this suggestion, but regardless of whether it is, or is not, correct, it seems very likely that the parafoveal-on-foveal effect that was observed here arose because the word that caused it was line final.

Before closing, there is a final aspect of the results that requires mention, namely, that response accuracy and RTs with respect to grammaticality decisions were comparable between sentences containing a transposition alone (92% accurate, 2261ms) and sentences containing an ungrammatical final word only (91% accurate, 2247ms). In the present experiment, when participants read sentences containing a transposed word pair, they detected that transposition equally as often, and as rapidly, as they detected a sentence-final word that was grammatically illegal. According to the OB1 Reader model, this pattern of effects should not occur because processing in accord with the spatiotopic map should ensure that word pairs appearing out of order are accommodated and interpreted comfortably and without difficulty. Thus, the degree to which readers detect ungrammaticalities that arise due to the transposition of a pair of words in the sentence should be substantially reduced relative to their detection of single ungrammatical words that have been substituted within the sentence. Clearly, this was not the case in the present study. It seems very likely that detection rates and RTs were comparable across these two conditions here because the present stimuli were well pre-screened and tightly controlled to ensure that both types of manipulation produced sentences that were seen as ungrammatical, and participants were not required to make speeded judgments. To be explicit, the present results demonstrate very clearly that there is nothing particularly special about transposed word pairs in sentences in relation to them signalling, or not signalling, ungrammaticality.

2.5.1. Limitations of the present study

As discussed previously (see Section 2.4.2), the present study contained 75% ungrammatical sentences while most previous research on word transpositions using the GDT task (e.g., Mirault et al., 2018, 2020) has employed a 50% ungrammatical to 50% grammatical sentences protocol. The imbalance of ungrammatical to grammatically correct sentences in the present study, could therefore have been a factor that contributed to the especially high accuracy rates across all four conditions. That said, some previous research has also utilised an unbalanced GDT. In particular, Huang and Staub (2021a: Experiment 1) used an equal number of grammatically correct sentences and sentences containing a word transposition. However, in their study, participants also read filler sentences which were rendered ungrammatical by removing one or more words in the sentence. Hence, in their study, Huang and Staub had a ratio of 31 ungrammatical to 21 grammatical sentences. Despite this imbalance, Huang and Staub did still obtain a *Transposed-Word* effect on accuracy in line with other studies. Consistent with this suggestion, in the present study, I also obtained a *Transposed-Word* effect in relation to accuracy replicating virtually all previous *Transposed-Word* effect studies in which sentences were presented naturally. Given both these pieces of evidence, even though the proportion of grammatical and ungrammatical sentences was not balanced in the present experiment, it seems unlikely that this influenced the pattern of effects.

A second point of note is that I was unable to conduct bimodal analyses of the eye movement data following the example of Huang and Staub (2021a). Recall that Huang and Staub were able to examine processing for sentences containing a word transposition that were judged as grammatical compared with sentences containing a word transposition that were judged as ungrammatical. This was not possible in the current experiment because there were insufficient numbers of errors for sentences containing a word transposition. That is, in the present study participants made very few errors in general. In fact, participants made on average approximately 2 errors for sentences containing a word transposition only. This meant that I could not explore whether the failure to detect a word transposition (as reflected in an incorrect grammaticality decision) also coincided with no sizeable disruption to processing upon fixating the critical ungrammatical target word. Hence, the present results cannot be used to directly distinguish between noisy-channel (e.g., Gibson et al., 2013) accounts and the account proposed by Huang and Staub (2021a). One final point to make is that, as discussed previously (see Section 1.7), the present results may not generalise beyond alphabetic spaced scripts, and in particular, English. Hence, further research is necessary to examine the effects of word transpositions which constitute ungrammaticalities in unspaced alphabetic (e.g., Thai) and logographic scripts (e.g., Chinese).

2.6. Conclusions

In summary, in the present experiment I manipulated whether 5-word sentences did, or did not, contain a transposed word pair and did, or did not, contain a grammatically illegal final word. I measured participants' eye movements as they read and made grammaticality decisions to these stimuli. Decision accuracy was high, and RTs were relatively long probably due to task requirements (reading for comprehension and grammaticality decision) and tightly controlled experimental stimuli. Importantly, no disruption to processing prior to readers fixating the first word of the transposed word pair was observed. Transposed words caused significant and rapid disruption to processing and an ungrammatical sentence final word attracted readers' fixations and caused increased re-reading. Overall, the results (arguably) are best accounted for by a serial eye movement model, though it is fair to say that none of the models of eye movement control adequately explains all aspects of the results. The patterns of effects obtained in this experiment, in the main, do seem to be consistent with those reported by Huang and Staub (2021a), and it is clearly the case that readers do detect transposed word pairs during reading when transpositions result in ungrammaticalities and that these are detected incrementally, word by word, as sentences are read.

Chapter 3. Parafoveal versus foveal word order violation effects in sentence reading

3.1. Introduction

Parafoveal processing is essential for skilled natural reading (Blythe, 2014; Tiffin-Richards & Schroeder, 2015; see Section 1.5). By partially pre-processing the upcoming word in the parafovea, readers reduce the time they need to process that word once they fixate it. Furthermore, parafoveal preprocessing enables readers to accurately target a saccade to the next word they aim to process. The boundary paradigm (Rayner, 1975; see Section 1.5.2) has been one of the primary experimental protocols used to investigate what information is extracted from the upcoming word(s) in parafoveal vision. Research utilising the boundary paradigm manipulates how one or more target words are presented in parafoveal vision. The parafoveal preview readers receive prior to fixating a target word could be identical to the target word or it could be different (invalid). The degree of mismatch between preview and target can be varied to examine whether readers are sensitive to different types of information parafoveally. For example, Vasilev and Angele (2017) showed in their meta-analysis of boundary paradigm studies that readers spend longer viewing a target word when its parafoveal preview consists of a string of letters which do not form a word than when it is an unrelated word. The authors interpreted this finding as evidence that the more *word-like* a preview is, the less it would disrupt processing at the target word since an unrelated word is still a valid word that exists in the language and hence may be pre-processed to a greater extent than a string of letters which do not form a word.

Over the last half a century multiple boundary paradigm studies have consistently shown that readers extract visual, orthographic, and phonological information from the upcoming word in the parafovea (Cutter et al., 2015; Rayner, 2009; Schotter et al., 2012; see Section 1.5.2). Furthermore, Vasilev and Angele (2017) showed that the cost to processing associated with a preview change is modulated by the type of invalid preview used. The more a preview mismatches the target, the more disruption to processing there is once the target word is fixated in foveal vision. This is known as a parafoveal preview effect. Importantly for the present experiment, boundary paradigm studies have typically investigated this parafoveal preview effect in the context of grammatically correct sentences but not when sentences contained a grammatical violation that was visible in foveal vision. Therefore, it is unclear whether a preview change would result in a similar processing cost when a sentence is ungrammatical. This is one issue that will be examined in the present experiment.

A major point of contention among eye-tracking reading researchers has been whether semantic and/ or syntactic information may be extracted parafoveally (see Andrews & Veldre, 2019; Schotter et al., 2012; see Section 1.5.2). As discussed previously, when Schotter et al. published their review in 2012, there were no existing studies showing semantic or syntactic parafoveal preview effects in English. Since then, several studies (e.g., Schotter, 2013; Schotter et al., 2015; Schotter & Jia, 2016; Veldre & Andrews, 2016a, 2016b, 2017, 2018a, 2018c) have documented evidence that semantic information may be processed parafoveally and can produce semantic parafoveal preview effects. These findings are important to the ongoing parallel versus serial lexical processing debate in the eye-tracking reading literature (see Zang, 2019) especially with respect to existing computational models of oculomotor control during reading.

3.1.1. Parallel and Serial processing computational models of reading

The most developed computational model that assumes serial lexical processing is E-Z Reader (Reichle et al., 1998; Reichle 2011), and more recently, its successor Über Reader (Reichle, 2021; Veldre et al., 2020). A detailed discussion of both models is available in Section 1.6.1. In summary, both models postulate that readers allocate attention to one word at a time and hence lexically identify words serially in an incremental fashion. Furthermore, according to the models, processing occurs in stages and proceeds over at least two distinct processing cycles such that readers begin linguistically processing a word in parafoveal vision and continue processing the same word in foveal vision until that word is lexically identified. Because of these assumptions, the classic E-Z Reader model would predict that there should be no lexical or post-lexical parafoveal preview influences at fixation, nor should there be any parafoveal-on-foveal effects. This is because a word can only be lexically identified once it has been brought into foveal vision, hence the semantic and syntactic information pertaining to that word may only be accessed once it is successfully identified.

Simulations of data using the baseline E-Z Reader model (e.g., Schotter et al., 2014) have, however, shown that semantic parafoveal preview effects may be possible even under strict serial processing assumptions. This is because the upcoming parafoveal word may vary in terms of how much processing it requires until it is lexically identified. Hence, the easier a word is to process, the faster readers can reach the point at which they are processing the semantic characteristics of the upcoming word in the parafovea. Crucially, the estimates derived from such simulations suggest that the probability for readers to obtain semantic information in parafoveal vision is relatively low (8% on average) meaning that the semantic parafoveal preview effects reported in the literature can still be harmonised within the serial framework of E-Z Reader.

In contrast to E-Z Reader and Über Reader, SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024) are computational frameworks based on Parallel Gradient attentional distribution principles (for a detailed discussion see Section 1.6.2). According to both models, attention is allocated to several words simultaneously during every fixation. Processing occurs in stages both in parafoveal and in foveal vision. There are two key distinctions that set SWIFT and SEAM apart from serial processing accounts. Firstly, since attention is allocated to multiple words simultaneously, lexical identification can occur without the need to fixate a word. Secondly, since multiple words are lexically processed in parallel, lexical, and post-lexical information pertaining to the upcoming word in the parafovea may influence processing on the fixated word (a semantic or syntactic parafoveal-on-foveal effect), or processing on the parafoveal word once it is fixated (a semantic or syntactic parafoveal preview effect). In contrast to serial models, SWIFT and SEAM, therefore, predict that such higher-order effects associated with parafoveal processing should be quite prevalent and robust.

OB1 Reader (Snell, van Leipsig, et al., 2018) provides an alternative parallel processing account (see Section 1.6.3). According to the model, lexical processing is parallel, however, lexical, and post-lexical information are not integrated across words during sentence reading. Consequently, the model postulates that there should be no higher-order (semantic or syntactic) parafoveal-on-foveal effects during sentence reading. In addition, words should be assigned to their corresponding sentential locations marked in a sentence-level representation. This sentence-level representation is guided by information regarding the length of each word and syntactic expectations which are updated as more words are successfully assigned into a sentential position. One of the key hypotheses that arises from the implementation of such a word position coding scheme is that when readers encounter words out of order, their moment-to-moment cognitive processing (as reflected by eye movements) should not be disrupted. Conversely, a reader’s sensitivity to detect the ungrammaticality created by such transposed word pairs should be reduced (or they might be completely insensitive to it) because the sentence-level representation may allow them to assign the words to their correct positions in text regardless of the position in which they were identified individually within the sentence.

3.1.2. Word transpositions

The first study to examine how word transpositions affect reading was conducted by Rayner, Angele, et al. (2013). Rayner, Angele and colleagues were interested in examining the impact of having a pair of target words transposed versus substituted by other words in parafoveal preview via the boundary paradigm (Rayner, 1975). The study had three conditions: identity preview (*white walls – white walls*), transposed-word preview (*walls white – white walls*); substitute word preview (*vodka clubs – white walls*). Importantly, the word transpositions did not constitute an ungrammaticality. Hence the two target words were grammatical both in preview and in foveal vision. Reading times on the region consisting of both target words were longer for the transposed-word preview than the identity preview. Processing was further disrupted for the substitute word preview especially for GPT and ROut. Rayner, Angele, et al. interpreted these findings as evidence that processing on a target region is generally disrupted when there is a preview change, and that disruption is reduced when preview and target have a larger overlap. In other words, some useful information could still be extracted from the transposed-word preview since both words were still visible in foveal vision (albeit in a different order) when they were fixated, while in the substitute preview condition there was minimal orthographic and lexical overlap between preview and targets.

Multiple empirical studies over the past six years have examined the impact of word transpositions on decisions about whether a sentence is, or is not, grammatical via the GDT (see Section 1.8). Overall, the consensus in the literature is that readers fail to detect such word transpositions significantly more often when a transposition appears in an otherwise grammatical sentence versus when one or more of the other words in the sentence are also ungrammatical (Hossain & White, 2023; Liu et al., 2020, 2021, 2022, 2024; Milledge et al., 2023; Mirault et al., 2018, 2020; Mirault, Vandendaele, et al., 2022, 2023; Snell & Grainger, 2019b; Snell & Melo, 2024; Spinelli et al., 2024; Wen et al., 2021a, b, 2024). This was initially termed the *Transposed-Word* effect (Mirault et al., 2018) and was taken as evidence that word position coding is noisy and lexical processing is parallel in line with the OB1 Reader model (Snell, van Leipsig et al., 2018). However, several recent findings seem to provide evidence that the *Transposed-Word* effect may be explained via alternative theoretical accounts (see Section 1.7) including the noisy channel account (Gibson et al., 2013) and a modified version of the E-Z Reader model (e.g., Huang & Staub, 2021a).

While the *Transposed-Word* effect on grammaticality decisions has received significant empirical attention, the effect of word transpositions that are ungrammatical on eye movements during reading is relatively understudied. To date there are only two published studies to directly examine the issue (Huang & Staub, 2021a; Mirault et al., 2020). Since an in-depth discussion of these studies has already been provided in Chapter 1 and Chapter 2, here I will provide only a brief summary.

Mirault and colleagues found no difference in reading times on the two transposed words, which they interpreted as evidence in support of OB1 Reader (Snell, van Leipsig, et al., 2018). Crucially, they only ever compared reading times on words which were always transposed and hence were always ungrammatical. Hence their finding may not necessarily elucidate whether word transpositions, per se, disrupt reading as reflected by comparisons of local eye movement measures for transposed versus non-transposed words. In contrast, Huang and Staub (2021a) did find significant disruption to reading on transposed words such that reading times were only inflated when readers detected the ungrammaticality as reflected by their response on the GDT. The results obtained by Huang and Staub are, thus, incompatible with OB1 Reader since any disruption associated with detecting a word transposition should be observed in global measures (e.g., accuracy) but not on any measure of early processing associated with a particular word. Both studies, however, contained word transpositions which were visible in both parafoveal and foveal vision. Hence, it remains unclear, whether readers may be sensitive to an ungrammatical word transposition in parafoveal preview and how such a sensitivity may potentially modulate the *Transposed-Word* effect.

3.2. The current experiment

The present experiment aimed to answer two key questions. Firstly, what, if any, is the role of parafoveal processing in the detection of word transpositions as reflected by accuracy on the GDT. The second question was how the relationship between the order of two target words in parafoveal versus foveal vision may affect moment-to-moment cognition as reflected by local eye movement measures on the word preceding the violation as well as the two target words that sometimes constituted the grammaticality violation. Both questions were investigated in relation to existing computational models of eye movement control during reading, those being E-Z Reader (Reichle et al., 1998; Reichle, 2011) and Über Reader (Reichle, 2021; Veldre et al., 2020), SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024), and OB1-Reader (Snell, van Leipsig, et al., 2018). This represents the first empirical investigation to my knowledge to independently manipulate the order of two target words in parafoveal and foveal vision via the boundary paradigm (Rayner, 1975) such that when the two words are presented out of order, they constitute a grammaticality violation (see Figure 3.1).

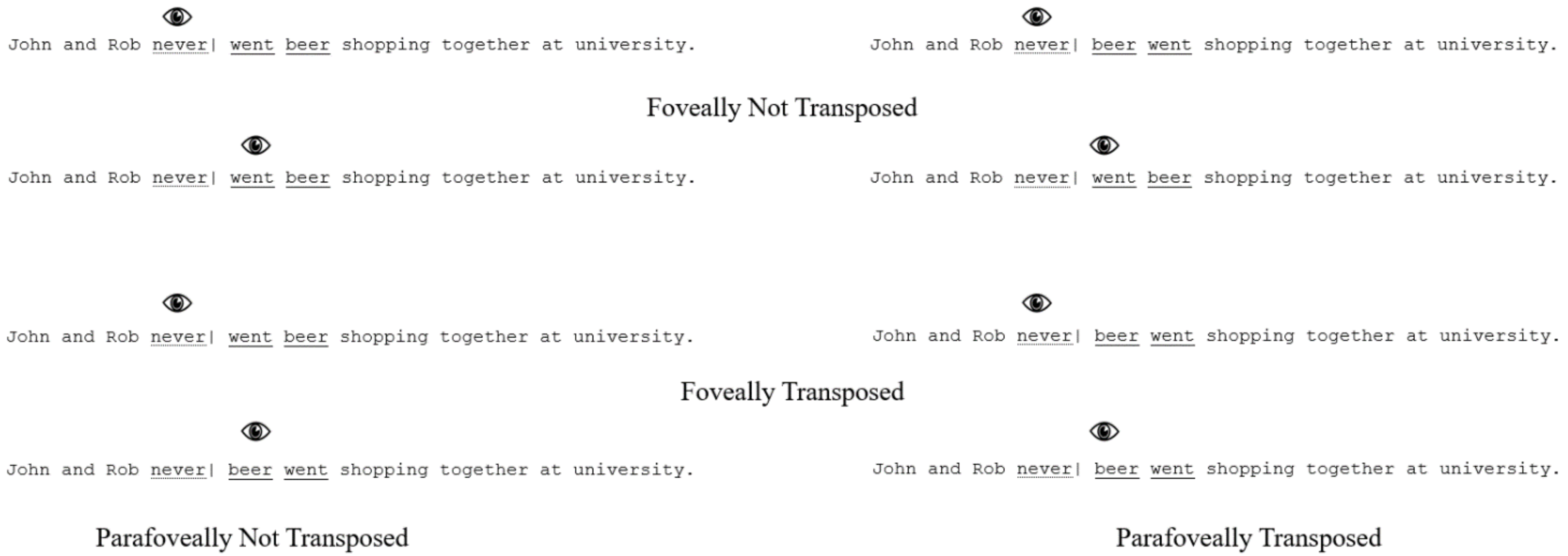


Figure 3.1. An example of the stimuli used in Experiment 2.

The pre-target word is underlined with a dotted line while both target words (target word 1 and target word 2) are underlined with a single line. None of the words and sentences were underlined within the actual experiment. The first sentence of each pair represents the parafoveal preview of the target words prior to the eyes crossing the invisible boundary. The second sentence of each pair represents how the two target words were presented in foveal vision. The invisible boundary is represented by the vertical line at the end of the pre-target word. The point of fixation for each sentence is indicated by the “👁” symbol.

Note that the Pre-Target region of interest was always occupied by the same (pre-target) word. Conversely, the Target Word 1 region of interest could be occupied by either target word 1 or target word 2 depending on whether the two target words were not transposed or transposed. Similarly, the Target Word 2 region of interest could be occupied by target word 2 (when the two targets were not transposed) or by target word 1 (when the two targets were transposed).

Previous research on word transpositions (e.g., Huang & Staub, 2021a; Mirault et al., 2018) has clearly shown that readers are better at judging grammatically correct sentences as such than judging sentences containing a foveal transposition as ungrammatical. Based on this, I would expect that accuracy on the GDT should be lower when the two target words are foveally transposed than when they are presented in their correct order in foveal vision. Since there is no previous research on the detection of transpositions in parafoveal vision, I can make no strong predictions based on experimental evidence. However, on the basis of the OB1 Reader (Snell, van Leipsig, et al., 2018) the presence or absence of a word transposition in parafoveal vision should not affect performance on the GDT since it is irrelevant whether the transposition is detected parafoveally or foveally, as long as a legal syntactic sentential representation can be created on the basis of the identified words.

Predictions regarding eye movement measures were driven by both existing empirical studies and the aforementioned computational models. As per E-Z Reader (Reichle, 2011) and Über Reader (Reichle, 2021; Veldre et al., 2020), there should be a significant effect of foveal transposition starting at the Target Word 1 region when the two target words were presented out of order, with longer first-pass reading times when the targets were transposed than when they were presented in the correct order. This prediction is also supported by the findings of Huang and Staub (2021a) who reported longer first-pass reading times on the target words when they were transposed versus when they were not. Although, I do note that in the case of Huang and Staub, the point of ungrammaticality was the second of the two transposed words, while in the present experiment, the point of ungrammaticality was the first of the two target words when they were transposed. According to both E-Z Reader (and Über Reader) there should be no first-pass reading effects of parafoveal transposition on the Pre-Target region since the models assume there is little to no parafoveal processing of semantic/ syntactic information. Importantly, there should be an interaction between parafoveal and foveal transposition on the Target Word 1 and potentially Target Word 2 regions which can be interpreted as a preview change effect as per Li et al. (2015). That is to say, any disruption associated with the foveal word transposition may be modulated by the presence or absence of a parafoveal transposition. When the two targets are parafoveally and foveally not transposed, processing should not be disrupted at all. When the two targets are transposed parafoveally but not transposed foveally, this would correspond to a preview change and should disrupt processing due to the mismatch between the two targets in parafoveal versus foveal vision. When the two targets are transposed both parafoveally and foveally, there should be disruption to processing associated with the ungrammaticality created by the transposition. Finally, when the two targets are parafoveally not transposed but foveally transposed, there should be disruption associated with the ungrammaticality created by the foveal transposition. In addition, processing may be further disrupted by the invalid preview. Any potential effects of parafoveal transposition on Target Word 1 and/ or Target Word 2 should therefore, only be observable if there is a significant interaction.

If readers process words in parallel, as per SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024), there should be an effect of parafoveal transposition, such that disruption should occur as early as on Pre-Target At the point of first fixating on the Target Word 1 region there should be both significant effects of parafoveal and foveal transposition as well as a significant interaction between the parafoveal and foveal transposition factors. In contrast to the serial processing accounts, however, SWIFT and SEAM would explain the interaction as evidence for parafoveal (post) lexical processing modulating foveal processing.

Alternatively, based on OB1 Reader model (Snell, van Leipsig, et al., 2018), both the parafoveal and foveal transposition factors should have no effect on early eye movement measures on any of the three regions of interest (Pre-Target; Target Word 1; Target Word 2). Yet, there should be a significant interaction between the two factors on Target Word 1 and potentially Target Word 2 such that whenever participants receive an invalid preview, they should exhibit a parafoveal preview change cost. Importantly, that cost should not be modulated by whether the two target words are foveally transposed or not as it should be purely driven by the mismatch between preview and targets in the preview change compared to identical preview conditions.

3.3. Methods

3.3.1. Participants

Fifty native English speakers with no known reading impairments and normal or corrected-to-normal vision from the student and staff community at the University of Central Lancashire, were recruited for this experiment. Participants were recruited via SONA and social media posts to participate in the current eye-tracking experiment. All participants received £25 in Amazon vouchers or 30-course credits to take part. 6 participants were excluded from data analysis (due to low accuracy on the GDT < 80%; poor synchronisation quality between the eye-tracking and electroencephalogram system or technical issues) and the data of 44 participants (*M* = 21.89 years, *SD* = 4.02 years, *Range* = 18:30, Female = 35) was included in the final analyses.

3.3.2. Design

The study was conducted as a 2(parafoveal transposition: not transposed vs. transposed) × 2(foveal transposition: not transposed vs. transposed) within-subjects Latin-square experiment. With this design, each participant saw one of the four versions of each sentence, so each sentence appeared an equal number of times across participants and its four versions.

3.3.3. Apparatus

Viewing was binocular but only participants' right eye movements were recorded by an SR Research Eyelink 1000 Plus system with a sampling rate of 1000 Hz. Participants were seated in front of an LCD monitor with 1920 by 1080 FHD resolution and 240Hz refresh rate at 70cm viewing distance. Stimuli were presented with a horizontal offset of 780 pixels from the centre of the monitor in black on a grey background and written in monospaced Courier New font size 24 with 2.3 letters subtending 1° of visual angle. The experiment was designed and presented via Experiment Builder v2.3.38 (SR Research).

3.3.4. Materials

One hundred and forty sentences were initially created to account for removal of sentences following norming studies. Each sentence was comprised of between 9 and 13 words, had a pair of target words (target word 1 and target word 2) which were manipulated such that they could be presented in the non-transposed, correct order (not transposed) or have their order swapped (transposed), such that the sentence contained a word transposition and became ungrammatical at the point of presenting target word 2 in the position of target word 1.

The pre-target word as well as both target word 1 and target word 2 were selected in the same manner as in Experiment 1 (see Section 2.3.5). the pre-target word was between 4 and 7 letters long (*M* = 5.31; *SD* = 0.98) and always a different length than the two target words. In addition, the target words were matched for length, each being between 4 and 6 letters long (*M* = 4.56; *SD* = 0.77). Furthermore, the Log Zipf lexical frequencies (van Heuven et al., 2014) of the pre-target, target word 1, and target word 2 ($p \geq 0.175$; see Table 3.1) were also matched.

All sentences were rated for their naturalness in a norming study by 10 participants who were native English speakers with normal or corrected-to-normal vision, no known reading impairments and who had attained a minimum of GCSE qualification via Prolific Academic (*M* = 22.8 years, *SD* = 3.26, *Range* = 18-28, female = 4). Participants were provided with £7 reimbursement for completing the naturalness questionnaire directly via the Prolific Academic platform. The ratings were given for each sentence in each of its two versions (foveally not transposed and foveally transposed) from 1 (very unlikely to hear this sentence in an everyday conversation) to 7 (very likely to hear this sentence in an everyday conversation). The t-test ($t_{(236)} = 65.25$, $p < 0.001$) analysis in R v4.3.1 (R Core Team, 2023) showed that the foveally not transposed sentences were rated as significantly more natural (*M* = 6.09, *SD* = 0.42, *Range* = 4.80:6.80) than those sentences containing a foveal transposition (*M* = 2.17, *SD* = 0.50, *Range* = 1.20:3.30).

Furthermore, to ensure that for the two target words there was no significant difference in predictability, 10 additional native English speakers with normal or corrected-to-normal vision, no known reading impairments, and who had attained a minimum of GCSE qualification were recruited via Prolific Academic (*M* = 23.7 years, *SD* = 3.47, *Range* = 19:30, Female = 3) and were asked to take part in a close probability task that was split in two parts. Each participant was presented with the initial words of each sentence twice, once up to and including the pre-target word, and once up to and including the first target of each sentence and asked to continue the sentence with the first word that came to their mind. As each participant saw the initial context up to and including the pre-target twice, the cloze-probability survey included all sentences for two experiments such that the sentences for one experiment would act as fillers for the sentences of the other experiment. Furthermore, all sentence beginnings were first presented only up to and including the pre-target in the first part of the task and the beginnings of all sentences including the first target were presented in the second part of the task. Participants were compensated with £9 for completing both parts of the task directly on the Prolific Academic platform. The within-subjects t-test on the second norming study data showed that the difference in cloze probability for the two targets was not significant ($t_{(236)}$ = -1.90, *p* = .058). Moreover, both targets had on average, low cloze probability values suggesting they were overall unpredictable given previous sentential context (target word 1: *M* = 2%, *SD* = 5%, *Range* = 0%:30%; target word 2: *M* = 4%, *SD* = 9%, *Range* = 0%:40%).

Table 3.1.
Descriptive statistics on Frequency for the pre-target word, target word 1, and target word 2. Frequency data was obtained and calculated via the Log Zipf scale (van Heuven et al., 2014)

Word	Frequency		t-test results (df = 236)	
	Range	M (SD)	pre-target	target word 1
pre-target	4.08 - 6.83	5.29 (0.74)		
target word 1	2.73 - 6.90	5.43 (0.86)	-1.36 (0.175)	
target word 2	3.02 - 7.19	5.40 (0.91)	-1.02 (0.309)	0.28 (0.782)

Note: These are the statistics for the final set of 119 sentences that were used for statistical analyses

3.3.5. Procedure

At the start of the experimental session, each participant was presented with an information sheet and consent form. All participants gave written informed consent before proceeding with the rest of the study. Consequently, they were asked to fill the Edinburgh handedness inventory (Oldfield, 1971) to confirm that participants were right-handed (*M* = 89%, *SD* = 16%, *Range* = 50:100%). Following that, they were tested for visual acuity with the Landolt C test (Precision Vision, La Salle, United States), to ensure that participants had 20/20 vision at 4m viewing distance (*M* = 0.09 errors, *SD* = 0.29, *Range* = 0:1 errors). Subsequently, a three-point horizontal calibration procedure was performed, to check that participants' eyes could be accurately tracked prior to starting the cap set-up.

All participants who successfully passed the three tests participated in two experiments within the same experimental session. Upon completing the first experiment in the session, they were provided with a short break before starting the second experiment. This chapter focuses on the second experiment in the session[4]. A three-point horizontal calibration was carried out at the beginning of each block of sentences and whenever necessary with an average ≤ 0.20° and a maximum ≤ 1.00° error. Each trial started with a drift check 48 pixels (1 degree of visual angle) to the left of the first letter of the first word. In the case of an error above the 0.2-degree threshold, calibration was carried out again. Following the drift check, a fixation cross was presented in the same location as the drift check dot, and participants had to look at the cross for 500ms before the sentence would appear. If a participant moved their eyes and did not look at the cross for 500ms from the onset of its presentation, the trial was recycled, and the researcher proceeded to either retry the trial or carry out a recalibration. Participants read 8 practice trials followed by 4 blocks of 31 sentences. Each block started with a filler sentence which was excluded from analyses and contained 30 experimental sentences for a total of 4 filler sentences and 120 experimental sentences.

When participants finished reading each stimulus, they had to look at a fixation cross presented at 715 pixels horizontal offset from the centre on the right-hand side of the screen for 150ms. Following that, participants were asked to determine whether the sentence they had just read was grammatically correct or incorrect on a separate screen. They used a low latency response box with the left button as the *No* answer and the right button as the *Yes* answer. When participants completed the experiment, they were provided with a

---

[4] The first experiment in the experimental session is discussed in the next chapter (4). The decision to present participants with Experiment 3 first was due to the fact that Experiment 3 contained no ungrammatical sentences and no manipulations to how target words appeared in foveal vision. Conversely, in Experiment 2, half of the sentences contained a word transposition which rendered the sentence ungrammatical in foveal vision. In addition, participants were instructed to make grammaticality decisions for all sentences in Experiment 2 while they were instructed to answer comprehension questions for some of the stimuli in Experiment 3. Hence, participants were always presented with Experiment 3 (Chapter 4) first, as it only required reading for comprehension.

questionnaire to determine to what extent they were aware of any boundary changes across the two experiments of the session.

The experiment was approved by the Ethics Committee at the School of Psychology and Humanities at the University of Central Lancashire (Ethics Reference: SCIENCE 0150 Phase 2).

3.3.6. Data pre-processing

All eye-tracking data was extracted via Data Viewer v4.1.211 (SR Research) such that a period of interest was selected starting from the onset of presentation of the sentence and ending with the onset of the presentation of the question. Following that, the 4-stage fixation cleaning procedure was carried out with only stage 4 selected such that fixations shorter than 50ms and longer than 1000ms were excluded. Further pre-processing steps were carried out in R v4.3.1 (R Core Team, 2023). Fixations on each interest area were excluded when the display change occurred early, during a fixation on a word prior to the gaze-contingent boundary, and when the display change was late (i.e., when the display change took more than 15ms - more than four screen refreshes based on a 240Hz refresh rate monitor, after fixation onset on target word 1). Fixations with hooks, when the display change was triggered early by a saccade that temporarily crossed the invisible boundary to finally end to the left of the boundary were also removed prior to analyses.

Fixations on the pre-target and either of the two target words in which participants made a blink and/or a skip were also removed from analyses. Observations on target word 1 and target word 2 where the saccade amplitude of the initial saccade to the critical interest area were ≥ 2.5 SD in absolute values from the mean saccade amplitude were also removed from analyses. Only observations from trials on which there was a first-pass reading fixation on any of the critical regions (pre-target; target word 1; target word 2) were considered.

Lastly, only observations from trials on which a first-pass reading fixation on the Pre-Target preceded a first-pass reading fixation on either Target Word 1 or Target Word 2 were considered to ensure that participants had fixated the pre-boundary word prior to crossing the boundary. These additional data pre-processing steps were undertaken because the experiment was run as a co-registered eye movement and electroencephalography study. Furthermore, only observations that could be matched between the eye movement and electroencephalography records following the procedure outlined by Degno et al. (2021) were considered, thus allowing me to directly compare any findings across the two data streams.

3.4. Results

3.4.1. Display Change Awareness

Out of the 44 participants, 30 participants reported that they had noticed one or more changes in the stimulus display throughout the experiment. One out of the thirty participants was unable to provide an estimate of how many times they detected a change. On average, the other 29 participants detected 21 changes (*M* = 20.83; *SD* = 27.70). A display change occurred on half of the experimental trials meaning that the order of the two target words was switched as the eyes crossed the invisible boundary in sixty of the experimental trials. This meant that participants were provided with an invalid parafoveal preview of the two target words for half of all trials. This is consistent with previous literature which has shown that readers do sometimes detect a display change (Degno et al., 2019a, 2019b; Slattery et al., 2011; Angele et al., 2016) in boundary paradigm studies.

3.4.2. Data analysis

Accuracy on the GDT was the only global measure analysed. Eight standard reading measures, as per Rayner (1998, 2009) were analysed for the Pre-Target, and the two target words regions (Target Word 1, and Target Word 2): FFD; SFD; GD; RP; GPT; ROut; TVT and RIn (see Section 1.3 and Section 2.4.1). [5]

For the Pre-Target, only the parafoveal transposition factor was used in the models for FFD, SFD, GD, GPT, RP and ROut as participants' eyes had not yet crossed the boundary at that point. For Accuracy, Pre-Target (TVT and RIn) as well as Target Word 1 and Target Word 2 models, both the parafoveal and foveal transposition factors were used. The contrast coding, model specifications and trimming procedures were identical to the one used for the corresponding measures in Experiment 1 (see Section 2.4.1).

3.4.3. Accuracy

With respect to accuracy, participants performed very well (*M* > 85% across all conditions, see Table 3.2) clearly showing that in most cases, the word transposition was detected and regarded as a syntactic violation. The effect of foveal transposition was significant and consistent with previous research (e.g., Huang & Staub, 2021a; Mirault et al., 2018) such that participants were better at judging foveally not transposed sentences as grammatically correct (a *Yes* response) versus judging foveally transposed sentences as ungrammatical (a *No* response).

There was no effect of parafoveal transposition and the interaction between parafoveal and foveal transposition approached significance. This interaction seemed to be driven by participants performing slightly worse (*M* = 86%) when the two targets were parafoveally not transposed but foveally transposed in comparison to when the two targets were both parafoveally and foveally transposed (*M* = 89%). On the other hand, performance was comparable when the two targets were foveally not transposed regardless of whether

---

[5] Skipping probability was not calculated because only observations on which a first-pass fixation was made for the corresponding region of interest (Pre-Target; Target Word 1; Target Word 2) were considered for analyses.

they were or were not transposed parafoveally (*M* = 96%). Therefore, it seems that readers were slightly better at judging sentences containing a foveal transposition as ungrammatical when they received a valid, in comparison to an invalid, preview while the relationship between preview and targets did not benefit or hinder their performance for grammatically correct sentences. It seems that the presence of a word transposition cannot be detected parafoveally but the detection of word transpositions in foveal vision can be improved when the transposition is visible both parafoveally and foveally (i.e., an identity preview). Additionally, the lack of an identity preview benefit when the two targets were foveally not transposed further shows that readers did not detect word transpositions parafoveally. Importantly, however, as discussed by Hossain and White (2023), Liu et al. (2022) as well as Milledge et al. (2023) the accuracy results alone, are not enough to confirm or falsify the predictions derived from OB1 Reader (Snell, van Leipsig, et al., 2018).

3.4.4. Eye movements

Next, let us focus on the eye-movement results for each of the three regions of interest, those being Pre-Target, Target Word 1, and Target Word 2 in turn. Additionally, all descriptive statistics are provided in Table 3.2. Further, in particular mean FFD, SFD and GD are visualised in Figures 3.2-4. All fixed effects estimates from the generalised linear mixed effects models can be found in Table 3.3.

With respect to Pre-Target, there was a significant effect of foveal transposition on TVT and RIn only such that participants were more likely to make a regressive saccade into the Pre-Target region and thus spent overall longer viewing it when the two target words were foveally transposed than not transposed. This effect was further qualified by a significant interaction. There was a clear gradation in both TVT and RIn such that they were lowest for sentences in which the two target words were both parafoveally and foveally not transposed. Both measures were inflated for sentences in which the two target words were parafoveally transposed but foveally not transposed. This likely reflects a general cost to processing associated with the preview change. When the two target words were parafoveally and foveally transposed TVT and RIn were substantially higher in line with participants detecting the ungrammaticality created by the transposition of the two target words in foveal vision and experiencing disruption associated with that ungrammaticality. Both TVT and RIn were further inflated when the two target words were parafoveally not transposed but foveally transposed. In other words, there were two distinct sources of disruption, namely the foveally transposed target words and the preview change. The presence of both a foveal transposition and a preview change resulted in more disruption to processing than the presence of either one on its own. In addition, there was no disruption to reading on any measure associated with the presence of a parafoveal transposition on its own and no evidence of early disruption to processing during first-pass reading of the Pre-Target (see Figure 3.2). These results suggest that there was little to no parafoveal sensitivity to the word transposition and fit with the lack of a parafoveal transposition effect on accuracy. Further Bayes factor calculations using the Dienes calculator (Dienes, 2008) adaptation in R (v 4.3.1) by Silvey et al. (2024) for FFD, SFD, GD, RP, ROut and GPT (see Table 3.3) provided confirmation that there were no significant differences in processing on the Pre-Target when the two target words were parafoveally transposed versus not transposed ($BF \leq 0.07$).

Importantly, both the serial processing frameworks of E-Z Reader (Reichle et al., 1998; Reichle, 2011) and Über Reader (Reichle, 2021; Veldre et al., 2020) as well as the parallel framework of OB1 Reader (Snell, van Leipsig, et al., 2018) can explain the absence of a syntactic parafoveal-on-foveal effect on Pre-Target during first-pass reading. In the first instance, if processing is serial as per E-Z Reader and Über Reader at the point of fixating the pre-target word during first-pass reading, semantic and syntactic information regarding target word 1 and/or target word 2 should not be accessible since word identification is strictly serial and sequential. Alternatively, as per OB1 Reader, while word identification can proceed in a parallel fashion, semantic and syntactic information pertaining to different words (e.g., pre-target, target word 1, and target word 2) is kept distinct and separate via the help of a spatiotopic sentence-level mechanism in working memory. Hence, while lexical information about target word 1 and target word 2 may already be available at the point of fixating the pre-target word during first-pass reading, the information pertaining to those words should be kept separate from the information pertaining to the pre-target word and therefore it should not interfere with its processing. Thus, according to this view, there should be no (post) lexical parafoveal-on-foveal effects on the Pre-Target. The lack of a parafoveal transposition effect on Pre-Target is potentially problematic, however, for the classic parallel processing frameworks of SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024), since the models posit parallel lexical identification without a mechanism for keeping the semantic and syntactic information pertaining to different words separate and distinct. Therefore, the SWIFT and SEAM models would predict a (post)lexical parafoveal-on-foveal effect on pre-target associated with the parafoveal transposition factor (although see also Schad & Engbert, 2012) which is clearly not visible in the current results.

The two late interactive effects on TVT and RIn, are potentially problematic only for OB1 Reader since they clearly show that local eye movement measures were inflated, when the two target words were foveally transposed than when they were not foveally transposed. Such results are difficult to fit within the OB1 Reader framework since the presence of a word transposition in foveal vision should produce effects in global (i.e., accuracy) but not local (e.g., TVT, RIn) measures.

With respect to Target Word 1, the effect of foveal transposition was significant for SFD, GD, GPT, TVT, ROut and RIn with longer fixation durations and a higher probability to make a regressive saccade into or out of Target Word 1 when the two target words were foveally transposed than not transposed. There was also a significant effect of parafoveal transposition on FFD, GD, GPT, ROut and RIn with longer reading times, as well as a higher probability to make a regressive saccade into and out of Target Word 1 when the two target words were parafoveally transposed than not transposed. All of these effects were qualified by significant interactions. First, with respect to FFD, SFD, and GD (see Figure 3.3), the presence of a preview change and the presence of a foveal transposition seemed to have a comparable effect on processing. Furthermore, having both a foveal transposition and a preview change (i.e., parafoveally not transposed but foveally transposed target words) did not result in additional disruption. With respect to ROut, GPT, TVT and RIn this trend changed. The disruption associated with having an invalid preview when the two targets were foveally not transposed, was diminished. Indeed, participants were equally likely to regress into Target Word 1 when the two targets were foveally not transposed regardless of whether they were parafoveally transposed or not transposed. Conversely, the disruption associated with the presence of a foveal transposition remained significant up to and including RIn. Importantly, having both an invalid preview and a foveal transposition became more disruptive than having a foveal transposition only, which was in turn more disruptive than having an invalid preview only for GPT, TVT, and RIn.

A potential reason for this divergence between the first-pass reading measures and later measures is that when the two targets were foveally transposed, Target Word 1 was ungrammatical. Hence it would require more processing than when it was grammatical. Moreover, when there was a preview change, Target Word 1 would presumably receive little to no parafoveal pre-processing. These two factors would mean that readers processed Target Word 1 to a lesser extent in parafoveal vision and required more processing time once they fixated it because it was ungrammatical. Such an explanation also fits with the patterns for

TVT and RIn observed for Pre-Target and further supports the notion that parafoveal processing has a limited role in the detection of word transpositions.

Overall, these results present a clear issue for OB1 Reader (Snell, van Leipsig, et al., 2018) since it is clear that the foveal transposition was detected and was disruptive to local first-pass reading eye movement measures. Additionally, the interactive patterns of effects on all measures (aside from RP) do provide further evidence that the disruption associated with the presence of a parafoveal transposition should be regarded as a general preview change cost rather than as sensitivity to the syntactic violation in the parafovea. The results on Target Word 1 do, however, fit within a classic parallel framework (SWIFT: Engbert et al., 2002; 2005, SEAM: Rabe et al., 2024) as well as a classic serial framework (E-Z Reader: Reichle et al., 1998; Reichle, 2011, Über Reader: Reichle, 2021; Veldre et al., 2020).

On Target Word 2, there was a significant effect of foveal transposition effect on all measures aside from FFD with longer reading times as well as a higher probability to refixate, regress out of and into Target Word 2 when the two targets were foveally transposed than not transposed. The same trend was also evident for FFD but failed to reach significance. These effects were reduced in comparison to the corresponding effects on target word 1 (see Table 3.2 and Figure 3.4) showing that the disruption associated with the foveal transposition was prolonged but not as strong as at the initial point of detection at Target Word 1.

In addition to the effects of foveal transposition, there was a significant parafoveal transposition effect only for GPT with longer GPT when the two targets were parafoveally not transposed than transposed. Importantly, besides the effects of foveal and parafoveal transposition, there was also a significant interaction between the two factors on GPT. There was also a significant interaction on TVT and ROut. For all three measures, the effect of parafoveal transposition was reduced when the two targets were foveally not transposed compared to when they were transposed. Consequently, all three interactive effects are consistent with and may be interpreted similarly to the patterns observed for GPT, TVT and RIn on Target Word 1.

Given all of this, the interactive effects across the three regions of interest collectively point to three conclusions. Firstly, the presence of a foveal transposition had a rapid, significant, and long-lasting disruptive effect on processing spanning both Target Word 1 and Target Word 2. Secondly, having an invalid preview when the two targets were foveally not transposed, conversely resulted in immediate but short-lived disruption to processing only on Target Word 1. Thirdly, when readers encountered a word transposition in foveal vision, they benefitted from having a valid preview, but that benefit was only visible in later measures following first-pass reading.

These results fit entirely with the serial processing accounts of E-Z Reader (Reichle et al., 1998; Reichle, 2011) and Über Reader (Reichle, 2021; Veldre et al., 2020). They are partially consistent with parallel accounts of SWIFT (Engbert et al., 2002; 2005) and SEAM (Rabe et al., 2024) since there were early significant effects of foveal transposition on both Target Word 1 and Target Word 2. However, the lack of any early effects on Pre-Target are inconsistent with SWIFT and SEAM since both models predict that the presence of a transposition should be detected parafoveally and produce disruption on the pre-target word. The lack of early effects on the Pre-Target is, conversely, compatible with OB1 Reader (Snell, van Leipsig, et al., 2018). The presence of early foveal transposition effects on Target Word 1 and Target Word 2 is, however, not compatible with the OB1 Reader model. Therefore, the results on all three regions of interest seem to primarily support a serial lexical processing account over any of the parallel processing accounts.

Table 3.2.

Descriptive statistics for all measures

Measure/ Region	Parafoveally and Foveally Not Transposed	Parafoveally and Foveally Transposed	Parafoveally Transposed and Foveally Not Transposed	Parafoveally Not Transposed and Foveally Transposed
Accuracy	96 (19)	89 (31)	96 (21)	86 (35)
Measure/ Region	Parafoveally Not Transposed		Parafoveally Transposed	
	Pre-Target			
FFD	232 (73)		233 (71)	
SFD	235 (72)		236 (69)	
GD	271 (105)		275 (108)	
RP	20 (40)		21 (41)	
GPT	302 (142)		305 (149)	
ROut	10 (30)		9 (28)	
Measure/ Region	Parafoveally and Foveally Not Transposed	Parafoveally and Foveally Transposed	Parafoveally Transposed and Foveally Not Transposed	Parafoveally Not Transposed and Foveally Transposed
TVT	346 (187)	460 (263)	378 (220)	489 (272)
RIn	15 (35)	45 (50)	24 (43)	53 (50)
	Parafoveally and Foveally Not Transposed	Parafoveally and Foveally Transposed	Parafoveally Transposed and Foveally Not Transposed	Parafoveally Not Transposed and Foveally Transposed
	Target Word 1			
FFD	254 (88)	277 (103)	278 (103)	267 (100)
SFD	257 (85)	307 (111)	299 (111)	316 (124)
GD	294 (124)	339 (154)	334 (136)	336 (167)
RP	18 (39)	23 (42)	24 (42)	24 (43)
GPT	324 (179)	439 (288)	392 (213)	420 (294)
ROut	7 (25)	20 (40)	14 (35)	19 (40)
TVT	379 (223)	619 (321)	445 (231)	681 (327)
RIn	18 (38)	48 (50)	18 (38)	58 (49)
	Target Word 2			
FFD	246 (81)	254 (95)	244 (81)	257 (102)
SFD	252 (75)	285 (98)	255 (87)	280 (99)
GD	279 (113)	301 (142)	281 (119)	299 (141)
RP	15 (36)	21 (41)	16 (37)	20 (40)
GPT	326 (182)	563 (236)	351 (236)	635 (425)
ROut	11 (32)	41 (49)	14 (35)	49 (50)
TVT	356 (186)	510 (261)	366 (189)	536 (262)
RIn	16 (36)	22 (42)	15 (36)	23 (42)

Note. Values for FFD, SFD, GD, GPT and TVT are given in milliseconds while values for accuracy, RP, ROut and RIn are given in percentages.

Table 3.3.

Fixed effects estimates from the generalised mixed effects models.

Measure/ Condition	Intercept				Parafoveal Transposition (Transposed versus Not Transposed)					Foveal Transposition (Transposed versus Not Transposed)					Interaction				
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
Accuracy	3.01	0.15	**19.96**	[2.71,3.30]	-0.02	0.16	-0.15	N/A	[-0.33,0.28]	-1.36	0.18	**-7.36**	N/A	[-1.72, -1.00]	0.67	0.28	2.43	N/A	[0.13,1.22]
Pre-Target																			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*		*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	236.53	5.35	**44.20**	[226.04,247.02]	0.61	3.66	0.17	0.02	[-6.57,7.79]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	243.13	7.22	**33.67**	[228.98,257.29]	-2.72	4.69	-0.58	0.02	[-11.91,6.46]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	276.88	6.77	**40.87**	[263.60,290.16]	4.00	4.52	0.88	0.02	[-4.87,12.86]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
RP	-1.52	0.11	**-13.32**	[-1.75, -1.30]	0.01	0.08	0.17	0.05	[-0.15,0.17]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GPT	310.00	7.99	**38.78**	[294.33,325.67]	2.76	5.46	0.51	0.02	[-7.95,13.47]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ROut	-2.49	0.12	**-19.98**	[-2.74, -2.25]	-0.08	0.15	-0.57	0.07	[-0.37,0.21]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	431.71	5.99	**72.07**	[419.97,443.45]	2.29	5.73	0.40	N/A	[-8.94,13.51]	121.75	5.62	**21.66**	N/A	[110.73,132.77]	-66.39	7.06	**-9.40**	N/A	[-80.24, -52.55]
RIn	-0.81	0.10	**-8.45**	[-1.00, -0.62]	0.12	0.08	1.49	N/A	[-0.04,0.28]	1.55	0.12	**13.31**	N/A	[1.32,1.78]	-0.99	0.16	**-6.35**	N/A	[-1.30, -0.69]
Target Word 1																			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	271.80	6.16	**44.16**	[289.42,332.21]	17.56	4.29	**4.10**	N/A	[-4.13,37.31]	3.69	**4.11**	0.90	N/A	[20.34,62.78]	-17.42	5.90	**-2.95**	N/A	[-28.97, -5.86]
SFD	310.82	10.92	**28.47**	[290.67,330.14]	16.59	10.57	1.57	N/A	[-2.29,31.45]	41.56	**10.83**	**3.84**	N/A	[25.88,62.17]	-54.13	21.08	**-2.57**	N/A	[-95.44, -12.81]
GD	327.81	8.77	**37.39**	[310.63,345.00]	22.98	5.82	**3.95**	N/A	[11.57,34.39]	23.53	**5.54**	**4.25**	N/A	[12.68,34.38]	-34.82	8.25	**-4.22**	N/A	[-50.99, -18.64]
RP	-1.42	0.10	**-13.60**	[-1.62, -1.21]	0.16	0.08	1.91	N/A	[0.00, 0.32]	0.17	0.08	2.07	N/A	[0.01,0.34]	-0.34	0.17	-2.05	N/A	[-0.67, -0.01]
GPT	395.66	6.91	**57.25**	[382.11,409.20]	49.49	6.69	**7.40**	N/A	[36.37,62.60]	70.47	7.46	**9.45**	N/A	[55.84,85.09]	-43.07	7.73	**-5.57**	N/A	[-58.22, -27.91]
ROut	-2.03	0.11	**-18.45**	[-2.24, -1.81]	0.41	0.13	**3.18**	N/A	[0.16,0.66]	0.84	0.12	**6.95**	N/A	[0.60,1.08]	-0.86	0.21	**-4.01**	N/A	[-1.28, -0.44]
TVT	549.82	6.82	**80.57**	[]536.44,563.19]	5.79	6.75	0.86	N/A	[-7.44,19.01]	254.27	6.15	**41.37**	N/A	[242.22,266.31]	-119.37	7.04	**-16.96**	N/A	[-133.16, -105.57]
RIn	-0.74	0.08	**-9.51**	[-0.89, -0.59]	-0.22	0.08	**-2.70**	N/A	[-0.37, -0.06]	1.73	0.11	**16.01**	N/A	[1.52,1.95]	-0.43	0.16	**-2.71**	N/A	[-0.75, -0.12]
Target Word 2																			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*		*Conf Int*	*β*	*SE*	*t(z)-value*		*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	252.09	4.91	**51.36**	[242.47,261.71]	-3.07	3.98	-0.77	N/A	[-10.88,4.73]	9.18	4.39	2.09	N/A	[0.57,17.79]	-6.00	6.30	-0.95	N/A	[-18.34,6.34]
SFD	274.26	7.18	**38.23**	[260.20,288.32]	5.70	6.42	0.89	N/A	[-6.89,18.29]	30.64	8.45	**3.63**	N/A	[14.08,47.19]	0.90	8.77	0.10	N/A	[-16.30,18.09]
GD	292.61	6.14	**47.67**	[280.58,304.64]	1.41	4.52	0.31	N/A	[-7.46,10.28]	20.32	5.58	**3.64**	N/A	[9.39,31.26]	-3.58	5.14	-0.70	N/A	[-13.65,6.49]
RP	-1.60	0.09	**-18.64**	[-1.77, -1.43]	0.02	0.10	0.22	N/A	[-0.18,0.22]	0.38	0.09	**4.28**	N/A	[0.21,0.56]	-0.01	0.18	-0.05	N/A	[-0.36,0.34]
GPT	474.73	6.84	**69.41**	[461.32,488.14]	-16.91	5.73	**-2.95**	N/A	[-28.13, -5.69]	270.41	6.20	**43.64**	N/A	[258.26,282.55]	-77.86	6.16	**-12.65**	N/A	[-89.92, -65.80]
ROut	-1.12	0.09	**-12.40**	[-1.30, -0.95]	-0.02	0.09	-0.21	N/A	[-0.21,0.17]	1.84	0.12	**15.04**	N/A	[1.60,2.07]	-0.67	0.22	**-3.00**	N/A	[-1.10, -0.23]
TVT	451.03	7.20	**62.62**	[436.92,465.15]	-8.70	7.28	-1.20	N/A	[-22.96,5.57]	168.23	8.09	**20.80**	N/A	[152.38,184.07]	-34.44	9.34	**-3.69**	N/A	[-52.74, -16.14]
RIn	-1.57	0.09	**-17.48**	[-1.75, -1.40]	-0.04	0.09	-0.42	N/A	[-0.21,0.14]	0.51	0.11	**4.47**	N/A	[0.29,0.74]	0.09	0.18	0.48	N/A	[-0.26,0.43]

Note. Significant terms are presented in bold, and terms approaching significance are underlined. All *p* values were adjusted via the Bonferroni correction with the summary(glht(<modelName>), test = adjusted("bonferoni")) function. N/A signifies that the factor or interaction was not considered in the model. *BF* refers to the Bayes factor calculated via an R adaptation of the Dienes calculator (Dienes, 2008) by Silvey et al. (2024).

^Confidence intervals (Conf Int) are calculated with the *confint* function in R with method = *Wald* at the 2.5 percentile and 97.5 percentile

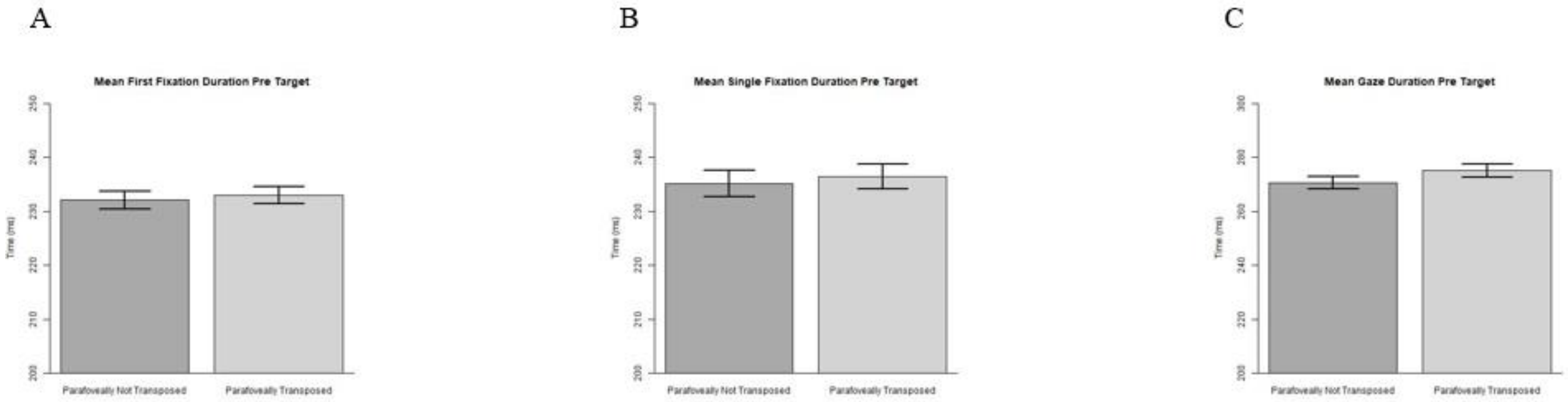


Figure 3.2. Mean first-pass reading times on Pre-Target

Panel A shows the mean first fixation duration on the Pre-Target region; Panel B shows the mean single fixation duration on the Pre-Target region; Panel C shows the mean gaze duration on the Pre-Target region.

Error bars represent standard errors.

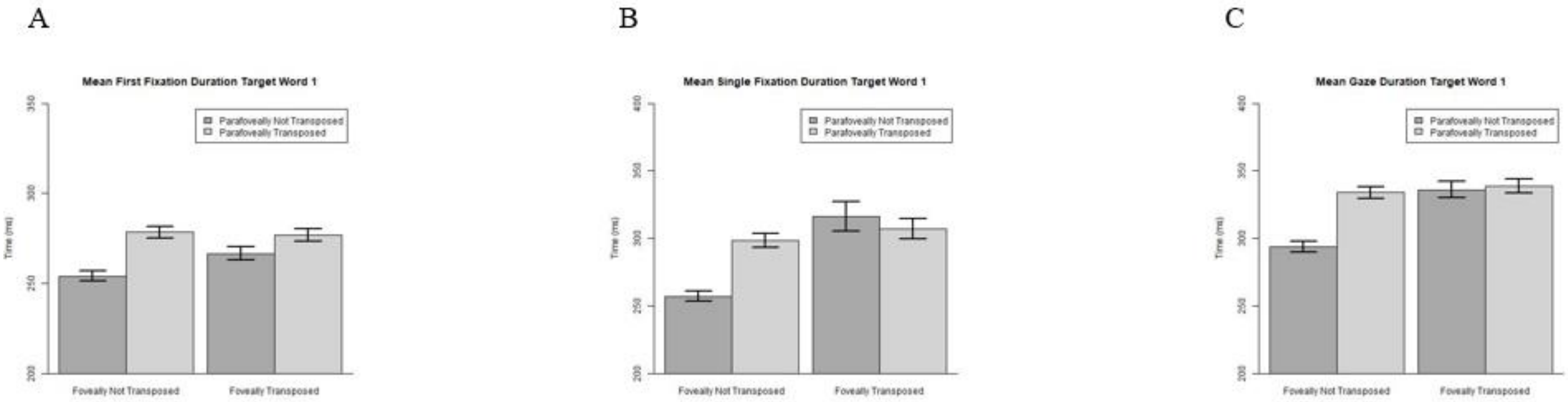


Figure 3.3. Mean first-pass reading times on Target Word 1

Panel A shows the mean first fixation duration on the Target Word 1region; Panel B shows the mean single fixation duration on the Target Word 1region; Panel C shows the mean gaze duration on the Target Word 1region.

Error bars represent standard errors.

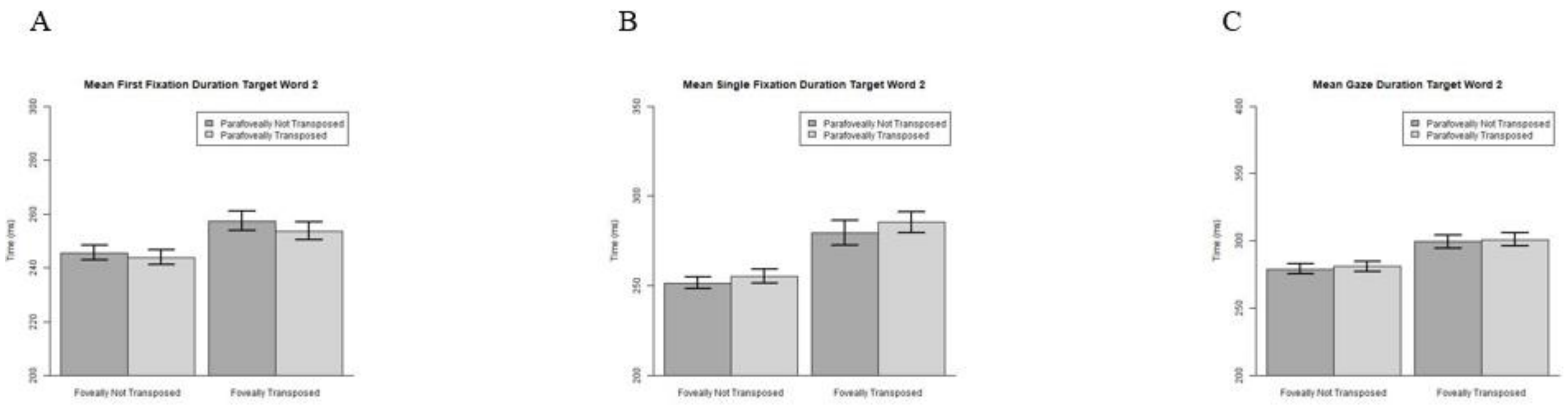


Figure 3.4. Mean first-pass reading times on Target Word 2

Panel A shows the mean first fixation duration on the Target Word 2 region; Panel B shows the mean single fixation duration on the Target Word 2 region; Panel C shows the mean gaze duration on the Target Word 2 region.

Error bars represent standard errors.

3.5. Discussion

The present experiment represents the first eye-tracking investigation of parafoveal processing in relation to the recently established *Transposed-Word* effect (Mirault et al., 2018; see Section 1.7; Section 3.1.2) under natural reading conditions. I orthogonally manipulated the order of two target words in parafoveal vision (not transposed versus transposed) and in foveal vision (not transposed versus transposed) via the boundary paradigm (Rayner, 1975). The two target words as well as the preceding (pre-target) word were carefully selected to ensure that any effects observed in the data were driven by the ungrammaticality created by transposing the two target words. I was particularly interested in whether readers are able to identify the word transposition parafoveally, before directly fixating either of the target words. Furthermore, I aimed to explore whether detecting a word transposition in parafoveal or foveal vision would disrupt reading and if that was indeed the case, when would that disruption become apparent in the eye movement record.

These questions are interesting with respect to the ongoing serial versus parallel processing debate in eye-tracking reading research because they provide a strong test of the parallel processing account of the recently established OB1 Reader model (Snell, van Leipsig, et al., 2018; see Section 1.6.3). In particular the model stipulates that word transpositions should be difficult to detect because sentential word order is managed by a spatiotopic sentence-level mechanism. Further, since this spatiotopic mechanism is independent from the mechanisms of oculomotor control, I expected that any disruption to processing associated with detecting a word transposition should only be observed on global (e.g., accuracy) but not local (e.g., FFD) measures. Furthermore, any disruption in the eye movement record should start at the point of fixating the first of the two target words and only occur when there is a preview change regardless of whether the two targets were foveally transposed or not.

The question of whether readers are able to parafoveally detect ungrammaticalities in general is relevant to the serial processing account of E-Z Reader (Reichle et al., 1998; Reichle, 2011) and its successor Über Reader (Reichle, 2021; see Section 1.6.1) as well as SWIFT (Engbert et al., 2002, 2005) and its successor SEAM (Rabe et al., 2024; see Section 1.6.2). This is because the two sets of models make diverging predictions with respect to how and when eye movement measures may or may not be affected by such parafoveal ungrammaticalities. If processing is indeed serial, I expected there to be no disruption on the pre-target word. Instead, disruption should occur at the point of fixating the first of the two target words and should be caused by either the preview change or the ungrammaticality created by the foveal word transposition. Alternatively, if processing is parallel as per SWIFT and SEAM, readers should have experienced disruption starting at the point of fixating the pre-target word associated with the presence of a word transposition in the parafovea. This disruption should continue when fixating the first of the two target words and should be modulated by whether there was or wasn’t a preview change.

Let us first turn to the results on accuracy and how they fitted against the predictions derived from the aforementioned models and existing literature. The first point of note with respect to accuracy is that readers performed substantially better when judging sentences in which the two target words were foveally not transposed than sentences which contained a foveal transposition. This finding is broadly consistent with existing literature (e.g., Huang & Staub, 2021a; Mirault et al., 2018, 2020) and fits with the predictions derived from the OB1 Reader model (Snell, van Leipsig, et al., 2018). The second point of interest is that the presence of a transposition in the parafovea did not impact accuracy suggesting that the presence of an ungrammaticality created by transposing two words is not detectable from the parafovea. Again, this is consistent with the predictions of the OB1 Reader model.

These findings can therefore be taken as evidence in favour of the processing account offered by OB1 Reader. Interestingly, in the present study stimuli were created in such a way as to maximise the chance for participants to fail to detect a word transposition as per the OB1 Reader model (see Section 3.3.4). Despite this, accuracy on the GDT for sentences containing a foveal word transposition was comparable to that reported in previous studies (e.g., Mirault et al., 2018, 2020). It is, therefore, unclear, why participants did not perform worse on the GDT in the present study compared to previous studies if *Transposed-Word* effects were driven by the processing mechanisms proposed by OB1 Reader. In addition to this issue, recent research (Hossain & White, 2023; Huang & Staub, 2023; Milledge et al., 2023; see also Mirault, Vandendaele, et al., 2022) shows that accuracy on the GDT for sentences containing a foveal word transposition is lower than for grammatically correct sentences even when sentences are presented serially word by word (see Section 1.7). Under such circumstances, readers presumable need to process each word serially and sequentially meaning that there is still a significant effect of word transpositions on accuracy even under conditions which enforce serial processing. Consequently, although the present findings on accuracy fit with the OB1 Reader model predictions, they cannot be taken as evidence in favour of that processing account.

Next, let us turn to the eye movement data and how it patterned in relation to the predictions derived from the five models discussed above. There were four key findings that I will discuss in turn. Firstly, there was no disruption to processing on Pre-Target associated with the presence of a parafoveal transposition. This lack of a (post)lexical parafoveal-on-foveal effect fits with the parallel processing account of OB1 Reader (Snell, van Leipsig, et al., 2018) as well as the serial accounts of E-Z Reader (Reichle et al., 1998; Reichle, 2011) and Über Reader (Reichle, 2021). However, it is problematic for the SWIFT (Engbert et al., 2002, 2005) and SEAM (Rabe et al., 2024) since both models would predict that the parafoveal transposition should have affected processing on Pre-Target.

Secondly, there was significant disruption to processing on Target Word 1. There were two distinct sources of disruption which affected both early and late measures. These sources of disruption were the presence of a foveal word transposition and the presence of a preview change. While the disruption associated with the preview change is consistent with all five computational accounts, the disruption associated with the foveal word transposition is particularly problematic for OB1 Reader and none of the other processing accounts. This is because, as mentioned above, according to the model, any disruption associated with a foveal transposition should only affect accuracy but not the local eye movement measures on Target Word 1.

Thirdly, having an identity preview when the two targets were foveally transposed resulted in no benefit to first-pass reading processing but instead benefitted later measures (e.g., GPT and TVT) suggesting that the classically observed identity preview benefit (see Rayner, 2009; Schotter et al., 2012) cannot be generalised to scenarios in which the target word is ungrammatical and can, therefore, not be integrated into context. This point is not necessarily problematic for any of the five discussed computational models of eye movements in reading. However, it does suggest that the identity preview benefit effect is not universal and can be delayed under circumstances such as the ones in the present experiment.

The fourth and final point is that at Target Word 2, processing was still disrupted, however, there was likely only one source of disruption, that being the presence of a word transposition in foveal vision. This is, again, only problematic for the OB1 Reader model for the same reasons as for the Target Word 1. Broadly, these four key findings further support the notion that word transpositions are not detected parafoveally and hence, parafoveal processing has little to no influence on whether word transpositions are detected or not which is consistent with the findings on accuracy. Conversely, word transpositions seem to be detected rapidly in foveal vision and cause significant and prolonged disruption to reading visible in early and late reading measures on both target words.

3.5.1. *Limitations of the present study*

While the present study does provide strong evidence that the detection of word transpositions which constitute an ungrammaticality occurs in foveal vision, there are some limitations with respect to how generalisable these results are, as well as, how informative they are with respect to the five models discussed in the thesis.

Firstly, similar to Experiment 1 (see Section 2.5.1), in the present study participants made very few errors on the GDT. Hence, it was impossible to conduct eye-movement analyses contingent on how participants judged sentence grammaticality. Given this, the results from the present study do not allow differentiation between noisy-channel accounts (e.g., Gibson et al., 2013) and post-lexical integration accounts (e.g., Huang & Staub, 2021a).

A further potential limitation in respect of the present experiment was that Target Word 2 actually fell in peripheral vision at the point of fixating the Pre-Target region. This is an important point because processing of information in the periphery is not as efficient as processing information in the fovea or parafovea. That said, while Target Word 2 may have been in peripheral vision, it still was within the perceptual span at the point of fixating on the Pre-Target region. Critically, when the two target words were transposed, the sentence became ungrammatical at the Target Word 1 region. Consequently, at the point of fixating on the Pre-Target, the upcoming word was itself ungrammatical.

The final key issue is that most previous research investigating word transposition effects has compared two types of ungrammatical sentences (e.g., Mirault et al., 2018, 2020).

The fact that a word transposition is harder to detect as an ungrammaticality on its own versus when there is an additional ungrammaticality presented in a sentence has been taken as evidence for parallel processing as per the OB1 Reader model (Snell, van Leipsig et al., 2018). Having both a word transposition and an ungrammatical final word versus only a word transposition in a sentence, however, is potentially problematic. This is because it is possible that having multiple ungrammaticality cues makes it easier to reach a grammaticality decision than having a single ungrammaticality cue. Hence, it may be beneficial for future investigations to compare word transpositions to other types of ungrammaticalities in isolation (see Chapter 2; Spinelli et al., 2024).

3.6. Conclusion

In sum, the present experiment provides clear evidence in favour of the serial processing accounts of E-Z Reader (Reichle et al., 1998; Reichle, 2011) and Über Reader (Reichle, 2021) over the parallel processing accounts of SWIFT (Engbert et al., 2002, 2005), SEAM (Rabe et al., 2024) and OB1 Reader (Snell, van Leipsig, et al., 2018). Furthermore, both the accuracy and eye movement results show for the first time that parafoveal processing has a minimal role in the detection of word transpositions and the disruption to processing associated with that detection.

Chapter 4. Effects of word order violations and letter masks in parafoveal vision during reading

4.1. Introduction

One of the most robust findings in eye-tracking reading research using the boundary paradigm (Rayner, 1975; see Section 1.5.2) from the past fifty years is that readers obtain a benefit to processing on a target word when they can parafoveally preprocess the target word while fixating the previous (pre-target) word (see Rayner, 2009; Schotter et al., 2012). In addition, some studies have shown that processing on the pre-target word may be affected by the orthographic (and, arguably, some other) characteristics of the upcoming target word in preview (Angele et al., 2008; Inhoff et al., 2000; Rayner, 1975; Starr & Inhoff, 2004).

A major issue of contention, however, has been whether readers are able to parafoveally process not only word *n+1* but also the following word *n+2*. If readers are able to parafoveally process word *n+2* then it follows that they may experience some parafoveal-on-foveal effect on either the fixated word *n* or the following word *n+1* and/ or a parafoveal preview effect upon fixating word *n+2*. The existence of such effects is important in respect of claims in the ongoing parallel versus serial lexical processing debate in the eye movement reading literature (see Vasilev & Angele, 2017). Such effects are also critical to the evaluation of computational models of oculomotor control during reading.

4.1.1. Computational models of oculomotor control during reading

The first two computational models I will focus on are E-Z Reader (Reichle et al., 1998; Reichle, 2011) and its successor, Über Reader (Reichle, 2021). A detailed discussion of the assumptions and principles of both models is provided in Section 1.6.1. Therefore, I will only focus on how parafoveal-on-foveal and parafoveal preview n+2 effects may be explained within the framework of both models. According to E-Z Reader (and Über Reader) words are processed serially and sequentially, such that once a word is identified, attention shifts to the next word so that linguistic processing of that word can begin. In addition, the model postulates that the visual characteristics of all visible words are processed simultaneously with words further from the point of fixation receiving less processing. Hence, from such a serial perspective, there are two possible explanations for the existence of parafoveal preview n+2 effects (see Section 1.5.2). Namely, readers may experience some disruption to processing purely because of the visual mismatch between word *n+2* in preview versus foveal vision. Alternatively, as shown by Schotter et al. (2014) it is theoretically possible for readers to identify word *n* and word *n+1* rapidly enough that parafoveal processing of word *n+2* can begin at the point of fixating word *n*. Notably, such effects should be rare, and in principle, should not be driven by the lexical characteristics of the n+2 preview. Therefore, parafoveal n+2 effects should be rare, smaller than parafoveal preview n+1 effects (see Vasilev & Angele, 2017) and may be visual or orthographic but not semantic or syntactic. Furthermore, any processing of word *n+2* should have no impact on the processing of either word *n* or word *n+1*, meaning that there should be no parafoveal-on-foveal n+2 effects under natural reading circumstances.

The second point of discussion are the perspectives offered by the parallel processing accounts of SWIFT (Engert et al., 2002, 2005) and its successor model – SEAM (Rabe et al., 2024), as well as the OB1 Reader (Snell, van Leipsig, et al., 2018) model. An in-depth discussion of SWIFT and SEAM is provided in Section 1.6.2 while OB1 Reader is discussed in Section 1.6.3. Similarly to the discussion for E-Z Reader and Über Reader, I will only focus on how parafoveal-on-foveal and parafoveal preview n+2 effects may be explained within these parallel processing frameworks. All three models stipulate that during a fixation, readers process the fixated word and the upcoming two words in the parafovea. Therefore, both parafoveal-on-foveal and parafoveal preview n+2 should be quite common and (potentially) linguistically driven. SWIFT and SEAM further postulate that such effects may be lexical since according to both models, semantic information is integrated across words. Conversely, OB1 Reader would predict no semantic parafoveal-on-foveal effects since semantic information is not initially integrated across words according to that model.

4.1.2. Parafoveal-on-foveal and parafoveal preview n + 2 effects

There has been mixed empirical evidence with respect to both parafoveal-on-foveal and parafoveal preview n+2 effects in the literature (see Vasilev & Angele, 2017; Section 1.5.2). Some studies (Angele et al, 2008, 2011; Rayner et al., 2007; Yang et al., 2010) have failed to find any effect associated with the parafoveal preview change of word *n+2* on any of the three words for which processing should be affected (word *n*; *n+1* and word *n+2*). Conversely, Kliegl et al. (2007) did find that readers were sensitive to the characteristics of word *n+2* in parafoveal vision, however that sensitivity was observed as a parafoveal-on-foveal effect on word *n* and word *n+1* instead of as a parafoveal preview effect on word *n+2*. Subsequently, several studies (Radach et al., 2013; Risse & Kliegl, 2012; Yang et al., 2009; Yan et al., 2010) further showed evidence for statistically significant but relatively small parafoveal-on-foveal and parafoveal preview n+2 effects, suggesting that because the effects are quite small, it is possible that non-significant findings may have arisen due to insufficient statistical power rather than a true null effect.

Vasilev and Angele (2017) further explored the existence of parafoveal preview n+2 effects in their meta-analysis of eleven experiments across nine studies. The authors found evidence for a small difference in processing on word *n+2* when its preview at the point of fixating word *n* was invalid versus when it was valid meaning that readers likely parafoveally process word *n+2* at the point of fixating word *n* to a very limited extent. As discussed previously (see Section 1.5.2) these effects can be explained by both the serial processing frameworks of E-Z Reader (Reichle, 2011; see Schotter et al., 2014), and Über Reader (Reichle, 2021; see Section 1.6.1), as well as the parallel gradient processing frameworks of SWIFT (Engbert et al., 2002, 2005), SEAM (Rabe et al., 2024; see Section 1.6.2), and OB1 Reader (Snell, van Leipsig, et al., 2018; see Section 1.6.3). Therefore, existing evidence regarding parafoveal preview n+2 effects is inconclusive with respect to whether readers lexically process words in a serial and sequential manner (as per E-Z Reader and Über Reader) or in parallel (as per SWIFT and SEAM, or OB1 Reader).

4.1.3. Word transpositions

One of the key predictions of the OB1 Reader model (Snell, van Leipsig, et al., 2018; see Section 1.6.3) is that readers should fail to detect an ungrammaticality created by transposing two words within a sentence which is otherwise grammatically correct. According to OB1 Reader, words are assigned to sentence positions on the basis of word length cues and syntactic expectations. This is achieved via a sentence-level spatiotopic mechanism such that readers rapidly form expectations regarding the length and syntactic category of words at specific positions. Furthermore, any disruption associated with detecting a word transposition should be seen in global measures such as accuracy or RTs on the GDT. Conversely, there should be little to no disruption associated with detecting a word transposition in terms of local eye movement measures prior to or at the point of encountering the transposition. Therefore, examining word transpositions has become a primary way to investigate the predictions derived from OB1 Reader. Furthermore, similarly to parafoveal-on-foveal and parafoveal preview n+2 effects, word transpositions have been utilised to examine whether lexical processing is serial or parallel.

In line with the predictions of OB1 Reader, Mirault et al. (2018) found that participants make more errors and take longer to decide whether a sentence is ungrammatical when it contains a word transposition only (e.g., *The white was cat big*) in comparison to a sentence containing both a word transposition and an additional ungrammaticality such as a syntactically illegal final word (e.g., *The white was cat slowly*). This *Transposed-Word* effect has been replicated in twenty two published studies to date (Hossain & White, 2023; Huang & Staub, 2021a, 2022, 2023; Liu et al., 2020, 2021, 2022, 2024; Milledge et al., 2023; Mirault et al., 2018, 2020; Mirault, Declerck et al., 2022; Mirault, Leflaec, et al., 2022; Mirault, Vandendaele, et al., 2022, 2023; Snell & Grainger, 2019b; Snell & Melo, 2024; Spinelli et al., 2024; Tiffin-Richards, 2024; Wen et al., 2021a, b, 2024).

Because a detailed discussion of the *Transposed-Word* effect is provided in Section 1.8, here I will only briefly outline the key findings and how they have been interpreted considering the debate on whether words are lexically processed in parallel or serially. As discussed above, Mirault et al. (2018, 2020) argued that this *Transposed-Word* effect is consistent with the parallel processing account of OB1 Reader (Snell, van Leipsig, et al., 2018). Liu et al. (2022), in contrast, argued that readers initially detect a word transposition, which causes disruption to processing. However, at a later stage, they are able to assign the transposed words into their correct positions. Liu and colleagues suggest that this is because the noisy bottom-up visual input indicating how the words are presented, is overwritten by the top-down expectations about how the words should be parsed to form a grammatically correct sentence, consistent with a *noisy-channel* processing account (e.g., Gibson et al., 2013).

Only two studies to date have examined the impact of ungrammatical word transpositions on eye movements during reading (Huang & Staub, 2021a; Mirault et al., 2020). Both studies have been discussed in depth in Section 1.7 as well as Section 2.1.3. Mirault et al. (2020) found no disruption to processing on either target word when the two targets were transposed and followed by a syntactically legal word compared with when they were transposed and followed by a syntactically illegal word. The authors suggested that these findings are consistent with OB1 Reader. In contrast, Huang and Staub (2021a) found significant and early disruption to reading on the target words when they were transposed versus when they were not transposed, but only when participants judged the sentences containing a word transposition as ungrammatical. Huang and Staub, therefore explained these findings as a consequence of an overlap between the integration of word *n* and word *n+1* into context consistent with a modified version of E-Z Reader (Reichle et al., 1998, 2011). If the integration of the two words into context overlaps, this would allow for top-down expectations regarding how the words should be parsed to rapidly overwrite the bottom-up visual input so that even though the words are presented out of order, they are parsed in their correct order resulting in no disruption to reading (no inflated reading times or fixation probabilities) and a failure to detect the ungrammaticality created by the word transposition reflected in an incorrect grammaticality decision.-On the other hand, when integration does not overlap, the transposition should be detected and immediately disrupt reading.

Overall, existing evidence on the *Transposed-Word* effect suggests that the failure to detect word transpositions is not necessarily a consequence of parallel lexical processing with noisy word position coding as assumed by OB1 Reader (see Hossain & White, 2023; Huang & Staub, 2021a, 2023; Liu et al., 2020, 2021, 2022; Milledge et al., 2023; however, see also Mirault, Vandendaele, et al., 2022; Snell & Melo, 2024). As discussed in Section 3.1.2 one issue that has not been investigated with respect to the *Transposed-Word* effect is how it may be modulated by parafoveal processing. In Experiment 2 (Chapter 3) I aimed to investigate how word transpositions affect processing in parafoveal versus foveal vision. A critical question that follows from Experiment 2 is whether, to what degree, and with what time course readers parafoveally process both the first target word (*n+1)* and the second target word (*n+2)* when they are transposed in preview.

4.2. Current experiment

Experiment 3 was aimed at examining two theoretical questions with respect to parafoveal processing during reading via the boundary paradigm (Rayner, 1975). Firstly, to what extent may processing be affected differently when the first word following an invisible boundary (target word 1) is masked by a string of letters that do not form a valid word versus when the second word following an invisible boundary (target word 2) is masked in preview (see Figure 4.1). That is to say, I wished to know how word *n+1* and word *n+2* are parafoveally processed in and of themselves, as well as in relation to each other during normal reading. This question has already been investigated to an extent in previous research (Angele et al., 2008; see Vasilev & Angele, 2017; see Section 1.5.2). However, there were a few methodological differences between the present study and the one conducted by Angele et al. (2008). Firstly, in their experiment Angele and colleagues manipulated the lexical frequency of the first word following the invisible boundary such that it could be either high or low. This was done to examine whether parafoveal processing of the second word following the boundary can be enhanced when the first word after the boundary requires less processing. In contrast to this, in the present experiment, I matched both target words for lexical frequency since they could be presented in their correct order or be transposed in preview and given the stipulations of the OB1 Reader model that a higher frequency upcoming word in the parafovea may inhibit processing of the fixated word (see Section 2.3.5). Furthermore, Angele et al. utilised letter masks which included vowels in their previews while in the present experiment letter mask previews consisted exclusively of consonants. This is potentially important because letter masks which include vowels may provide more meaningful phonological information parafoveally than letter masks which do not contain vowels since some combinations of vowels and consonants may form syllables. As discussed in Section 1.5.2 (see Ashby & Rayner, 2004: Experiment 2) readers may parafoveally process the phonological information pertaining to the upcoming word in the parafovea not only at the level of single letters, but also at the level of syllables. Therefore, it is possible that readers can parafoveally process letter masks containing both vowels and consonants to a greater extent than letter mask containing only consonants.

Despite these methodological differences, I do expect that any effects observed in relation to my first research question should be broadly consistent with the results obtained by Angele and colleagues. In particular, I expect that there should be no effect of masking either one or both target words in preview at the pre-target or second target. Additionally, I would expect significant disruption to processing at the first target when it was masked by a letter string in preview regardless of whether target word 2 was masked or not in preview. Furthermore, having a letter mask preview of target word 2 should have little to no effect for any measure on Target Word 1. The lack of any significant parafoveal-on-foveal or parafoveal preview n+2 effects is consistent with the serial processing assumptions of E-Z Reader (Reichle, 2011) and Über Reader (Reichle, 2021). However, I do note such results can also be explained by the processing accounts of SWIFT (Engbert, 2002, 2005; see Schad & Engbert, 2012), and therefore, SEAM (Rabe et al., 2024), as well as OB1 Reader (Snell, van Leipsig, et al., 2018). This is because, all three models suggest that while words are processed in parallel, the number of words which are attended to (and therefore processed) simultaneously, is adaptable. That is to say, when words are easier to lexically identify (require less processing), attention is allocated to more words and further from the point of fixation. In contrast, when readers struggle with lexical identification (see Section 1.6.3), attention is allocated to fewer words and closer to the point of fixation resulting in oculomotor patterns similar to the ones predicted by serial processing accounts such as E-Z

Reader and Über Reader. Consequently, the lack of parafoveal-on-foveal and parafoveal preview n+2 effects may be explained by the difficulty in processing word *n*, and word *n+1* resulting in attention not being allocated to word *n+2* at the point of fixating word *n*.

The second question of interest is whether there is any difference in the cost to processing associated with two target words being transposed, and hence ungrammatical in preview versus the two target words being masked by strings of letters which do not form a valid word in preview (see Figure 4.1). This question is particularly relevant to the OB1 Reader model (Snell, van Leipsig, et al., 2018) since word transpositions should not affect any local eye movement measures. Consequently, if processing is parallel and words are assigned to sentence positions based on length expectations as per OB1 Reader (see Section 1.6.3), then a transposed pair of target words in preview should disrupt processing substantially less than a masked pair of target words in preview. Any disruption associated with transposing or masking the two target words in preview should be driven by the visual and orthographic mismatch between how the target words are presented in parafoveal versus foveal vision. This mismatch should be significantly lower when the two Targets are transposed versus when they are masked by completely different letter strings.

If processing is serial as per E-Z Reader (Reichle, 2011) and Über Reader (Reichle, 2021), then there should be no disruption to processing on the pre-target and potentially second target regions. There should, however, be significant disruption at the Target Word 1 region associated with having a preview change. This is because at the point of fixating the Pre-Target, readers should also parafoveally process Target Word 1 (at least). Consequently, having an invalid preview of the Target Word 1 (either because of the transposition or because of the two target words being masked) should disrupt processing at the point of fixating the first target, consistent with a parafoveal preview change cost.

It is possible that disruption for target words that are masked might be comparable to disruption associated with a preview of transposed target words since both previews are invalid and should make parafoveal processing of Target Word 1 more difficult and less useful. Alternatively, a pair of transposed target words would provide readers with a valid word in preview allowing for deeper parafoveal processing of the preview. Conversely, a masked targets preview would provide readers with a string of letters that do not comprise a valid word meaning that they should not be able to parafoveally process the preview of Target Word 1 to the same extent as when the two target words are transposed in preview. Consequently, there may be more disruption at the Target Word 1 region when in preview the two target words were transposed than when they were masked. As noted previously, both questions were investigated in relation to existing computational models of eye movement control during reading, those being E-Z Reader (Reichle et al., 1998; Reichle, 2011), Über Reader (Reichle, 2021; Veldre et al., 2020), SWIFT (Engbert et al., 2002; 2005), SEAM (Rabe et al., 2024), and OB1-Reader (Snell, van Leipsig, et al., 2018).

Figure 4.1. An example of the stimuli used in Experiment 3.

The pre-target word is underlined with a dotted line while both target words (target word 1 and target word 2) are underlined with a single line. None of the words and sentences were underlined within the actual experiment. The first sentence of each pair represents the parafoveal preview of the target words prior to the eyes crossing the invisible boundary. The second sentence of each pair represents how the two target words were presented in foveal vision. The invisible boundary is represented by the vertical line at the end of the Pre-Target region. The point of fixation for each sentence is indicated by the “👁” symbol.

The Pre-Target region was always occupied by the same (pre-target) word. Conversely, the Target Word 1 region of interest could be occupied by target word 1, target word 2, or a letter mask depending on the preview change condition. Similarly, the Target Word 2 region of interest could be occupied by target word 2 (when the two targets were not transposed in preview), by target word 1 (when the two targets were transposed in preview), or by a letter mask in preview depending on the preview change condition.

4.3. Methods

4.3.1. Participants

I tested a total of 53 native English speakers with no known reading impairments and normal or corrected-to-normal vision from the student and staff of the University of Central Lancashire community, who did not take part in the norming studies. Participants were recruited via SONA and social media posts to participate in the current eye-tracking experiment. All participants received £25 in Amazon vouchers or 30-course credits to take part. 8 participants were excluded from data analysis due to poor synchronisation quality between the eye-tracking and electroencephalogram systems, low accuracy on the comprehension questions (<75%) or being ambidextrous, and the data of 45 participants (*M* = 21.69 years, *SD* = 3.77, Female = 36; Age Range = 18-30) was included in the final analyses.

4.3.2. Design

There were five experimental within-subjects conditions and in each condition, one or two target words were manipulated in parafoveal preview via the boundary paradigm (Rayner, 1975). These five conditions were: an identical preview for both target words (identity), a masked preview for target word 1 and an identical preview for target word 2 (target word 1 masked, target word 2 not masked), an identical preview for target word 1 and a masked preview for target word 2 (target word 1 not masked, target word 2 masked), a masked preview for target word 1 and target word 2 (target word 1 masked, target word 2 masked), a transposed-word preview for target word 1 and target word 2 (transposed target words). In the analyses, I grouped these conditions in two ways. The first grouping was a 2(target word 1 mask: masked versus not masked) by 2(target word 2 mask: masked versus not masked) design according to which, I considered all conditions aside from the transposed targets preview condition. The second grouping was of one factor with three levels (preview type: identity preview versus transposed targets preview versus masked targets preview).

4.3.3. Apparatus

Viewing was binocular but only participants' right eye movements were recorded by an SR Research Eyelink 1000 Plus system with a sampling rate of 1000 Hz. Participants were seated in front of an LCD monitor with 1920 by 1080 FHD resolution and 240Hz refresh rate at 70cm viewing distance. Stimuli were presented with a horizontal offset of 780 pixels from the centre of the monitor in black on a grey background and written in monospaced Courier font size 24 with 2.3 letters subtending 1° of visual angle. The experiment was designed and presented via Experiment Builder v2.3.38 (SR Research).

4.3.4. Materials

One hundred and seventy sentences were initially created to account for removal of sentences following norming studies. Each sentence was comprised of between 9 and 13 words and had a pair of target words (target word 1 and target word 2) that were always foveally presented in their correct order (see Figure 4.1). In preview, the two target words could have their order transposed resulting in an ungrammaticality. Furthermore, for each target word, a letter mask was created that matched the target for length, was comprised of consonants only, in combinations that did not form any lexical units (e.g., phonemes, syllables, or words), and in which each letter substitute was visually similar to the original letter it was used to replace. Either one or both targets could be replaced by their corresponding letter masks in preview (see Figure 4.1).

The pre-target word and the two target words were selected following the same procedures as for Experiment 1 and Experiment 2 (see Section 2.3.5). The pre-target word was between 4 and 7 letters long (M = 5.59, SD = 1.08) and always a different length than the target words. In addition, I matched the target words for length, each being between 4 and 6 letters long, and not different in any of the conditions (M = 4.81, SD = 0.82).

The length of the pre-target word was always different from the length of either of the two target words. Furthermore, the pre-target, first, and second target words were matched on their Log Zipf lexical frequency (van Heuven et al., 2014; see Table 4.1) to ensure they were comparable ($p \geq 0.095$).

Table 4.1.
Descriptive statistics on Frequency for the pre-target word, target word 1, and target word 2. Frequency data was obtained and calculated via the Log Zipf scale (van Heuven et al., 2014)

Word	Frequency		t-test results (df = 298)	
	Range	M (SD)	pre-target	target word 1
pre-target	4.11 - 6.83	5.21 (0.69)		
target word 1	2.23 – 7.19	5.37 (0.96)	-1.68 (0.095)	
target word 2	1.97 - 7.19	5.23 (0.94)	-0.26 (0.798)	1.26 (0.210)

Note: These are the statistics for the final set of 150 sentences that were used for statistical analyses

All sentences were rated for their naturalness in a norming study by a separate group of 10 participants who were native English speakers with normal or corrected-to-normal vision, no known reading impairments, and who had attained a minimum of GCSE qualification via Prolific Academic (*M* = 22.8 years, *SD* = 3.26, range = 18-28, female = 4). Participants received £7 reimbursement for completing the naturalness questionnaire directly via the Prolific Academic platform. The ratings (*M* = 5.97; *SD* = 0.4; *Range* = 4.90:6.80) were given for each sentence from 1 (very unlikely) to 7 (very likely to hear this sentence in an everyday conversation). Only sentences that were deemed natural (i.e., received ratings higher than 4) were used for the main experiment.

To ensure that the two target words did not differ in predictability, 10 additional native English speakers with normal or corrected-to-normal vision, no known reading impairments, and who had attained a minimum of GCSE qualification were recruited via Prolific Academic (*M* = 23.7 years, *SD* = 3.47, *Range* = 19:30, Female = 3). The participants were asked to take part in a close probability task that was split into two parts. Each participant was presented with the initial words of each sentence twice, once up to and including the pre-target word, and once up to and including the first target of each sentence, and asked to continue the sentence with the first word that came to their mind. As each participant saw the initial context up to and including the pre-target twice, the cloze-probability survey included all sentences for two experiments such that the sentences for one experiment acted as fillers for the sentences of the other experiment. Furthermore, all sentence beginnings were first presented only up to and including the pre-target in the first part of the task, and the beginnings of all sentences including the first target word were presented in the second part of the task. Participants were compensated with £9 for completing both parts of the task directly on the Prolific Academic platform. The within-subjects t-test on the second norming study data showed that the difference in cloze probability for the two Targets was not significant ($t_{(298)}$ = -0.07; *p* = .945) meaning that the two target words were comparable in terms of their contextual predictability. Moreover, both target words had low cloze probability values suggesting they were overall unpredictable (target word 1: *M* = 3%; *SD* = 9%; *Range* = 0%:40%; target word 2: *M* = 3%; *SD* = 8%; Range = 0%:40%).

4.3.5. Procedure

The procedure for Experiment 3 was identical to the one used for Experiment 2 (see Section 3.3.5) except for three distinctions. As for Experiment 2, prior to the start of the experimental session participants were asked to complete the Edinburgh handedness inventory (*M* = 90%; *SD* = 15%; *Range* = 50:100% right-handed) to confirm that they were right-handed. Following that, they were tested for visual acuity with the Landolt C test

(Precision Vision, La Salle, United States), to ensure that participants met 20/20 vision at 4m viewing distance (*M* = 0.09; *SD* = 0.29; *Range* = 0:1). Third, a three-point horizontal calibration procedure was performed, to check that participants' eyes could be accurately tracked prior to starting the EEG cap set-up.

The experiment itself was conducted broadly in the same manner as Experiment 2 (see Section 3.3.5). The first difference in comparison to Experiment 2 was in the number of sentences participants read. There were 10 practice trials followed by 5 blocks of 31 sentences. Each block started with a filler sentence which was excluded from analyses and contained 30 experimental sentences for a total of 5 filler sentences and 150 experimental sentences.

The second distinction from Experiment 2 was in the placement of the fixation cross to the right of the sentence. The fixation cross used to terminate the presentation of the sentence was presented at 775 pixels horizontal offset from the centre on the right-hand side of the screen. The third, and most important distinction was that in Experiment 3, participants were asked comprehension questions for some of the sentences. For two of the practice trials, two of the filler trials, and 38 of the experimental sentences, participants were required to answer a short reading comprehension question associated with the corresponding sentence. The question was presented on a separate screen following the end of the corresponding trial. Participants were instructed to use a response box with the left button as the *No* answer and the right button as the *Yes* answer.

4.3.6. Data pre-processing

I utilised an identical set of pre-processing procedures and steps in this experiment as in Experiment 2 (see Section 3.3.6) with one exception. Observations were not removed if participants answered incorrectly on the comprehension questions in Experiment 3, while in

Experiment 2, observations were removed if participants made an incorrect grammaticality decision.

4.4. Results

4.4.1. Comprehension accuracy

Accuracy on the comprehension questions was very high (*M* = 91%; *SD* = 5%; *Range* = 76:97%) across all participants who were included in the analyses. This pattern clearly shows that participants read and understood the stimuli they were provided with.[6]

4.4.2. Display Change Awareness

Out of the 45 participants, 29 participants reported that they had detected at least one change in the stimulus display. Out of the 29 participants, one was unable to provide an estimate of how many times they detected a change. On average, the other 28 participants detected 17 changes (*M* = 16.54; *SD* = 17.49). A display change occurred in four-fifths of the experimental trials meaning that participants received an invalid parafoveal preview of either the first, the second, or both target words following the boundary on 120 out of the 150 experimental trials. These results fit with previously reported detection rates of the display change in boundary paradigm studies (e.g., Degno et al., 2019a, 2019b; Slattery et al., 2011; Angele et al., 2016).

4.4.3. Data analysis

The descriptive statistics for each condition and each region of interest are provided in Table 4.2. Further, the means for FFD, SFD and GD across the five conditions are visualised in Figures 4.2-4. I conducted two sets of analyses on the eye movement data. The first set of

---

[6] Accuracy was measured at the beginning of data preprocessing and reflects performance for all the questions to experimental sentences prior to the removal of observations as per the data pre-processing procedures applied.

analyses was aimed at investigating the effect of letter masks on Target Word 1 versus Target Word 2 regions. For that set of analyses, I included both the target word 1 mask and the target word 2 mask factors which had two levels. The fixed-effects estimates from these models are provided in Table 4.3. The second set of analyses was aimed at exploring to what extent readers are sensitive to word transpositions in the parafovea. To do that, I ran models only considering the identity preview condition, the transposed targets preview condition, and the masked targets condition. This allowed me to compare processing when there was no change between preview and target words versus when the change in preview was created by transposing the two target words versus when there was a change created by masking both target words with strings of consonants that could not form a word in preview. The fixed-effects estimates from the second set of analyses are provided in Table 4.4. The model specifications and trimming procedures were identical to the ones used for Experiment 1 and Experiment 2 (see Section 2.4.1).

4.4.4. Interactive models

On Pre-Target, there were no reliable first pass effects (see Figure 4.2), but there was a significant effect of target word 1 being masked in the parafovea on TVT and RIn only. Readers were more likely to make a regressive saccade into the Pre-Target and hence spend overall longer reading it when target word 1 was parafoveally masked than when it was not masked. This is a relatively late effect which indicates that readers experienced processing difficulty after their eyes had crossed the invisible boundary when processing of Target Word 1 was disrupted due to it being masked by a string of letters in preview. There were no other significant effects or interactions on pre-target suggesting that when either one or both target words were masked in preview, this did not influence processing on the Pre-Target. To further confirm the null effects observed on the Pre-Target region, I utilised the Dienes calculator (Dienes, 2008) adaptation in R (v 4.3.1) by Silvey et al. (2024) for FFD, SFD, GD, RP, ROut and GPT (see Table 4.3). All Bayes Factors calculated for both the simple effects of each factor, as well as their interaction, showed no clear evidence in support of the alternative hypothesis ($BF \leq 0.38$) suggesting that reading times, refixation probability and regression out probability were comparable across the four types of sentences. This pattern fits with previous research on n+2 parafoveal-on-foveal effects (e.g., Rayner et al., 2007; see Vasilev & Angele, 2017). Moreover, the lack of any early disruption associated with having a letter string instead of target word 1 in parafoveal vision further supports the idea that the parafoveal processing of Target Word 1 was limited when it was masked and fits with the patterns observed for RIn and TVT. That is, the lack of early effects and the presence of late effects seem to point to processing difficulty arising due to the decreased parafoveal pre-processing of the first target word when it was masked versus when it was not in preview at the point of fixating Pre-Target.

On Target Word 1, reading times (see Figure 4.3) and fixation probabilities were significantly disrupted when the first target word was masked in preview. This trend was visible for all measures aside from RIn. These results are unsurprising and consistent with a strong n+1 preview effect which has been extensively documented in the literature (see Cutter et al., 2015; Vasilev & Angele, 2017). There was no significant disruption associated with having a letter string mask compared to an identity preview for the second target word. This was similar to the patterns observed at the Pre-Target region. Importantly, there were significant interactions between target word 1 mask and target word 2 mask on GD, GPT, and TVT. For all three measures, it was evident that having a masked preview of target word 2 was disruptive to reading but only when target word 1 was not masked in preview. All three measures were lowest for the identity preview condition consistent with an identity preview benefit. Next, there was an increase in all three measures when only target word 2 was masked in preview. Reading times were further substantially inflated and comparable when target word 1 was masked in preview regardless of whether target word 2 was also masked in preview or not. In other words, disruption to processing was comparable when only the first target word was masked in preview versus when both targets were masked. Taken together, these patterns indicate that when parafoveal processing of Target Word 1 was disrupted by its invalid preview, readers likely did not parafoveally process Target Word 2. Conversely, when target word 1 was not masked in preview, readers were potentially able to parafoveally pre-process the second target, albeit to a lesser degree than they were able to parafoveally process the first target word. Furthermore, FFD and SFD patterned numerically in the same way as GD, although the interaction effect was only statistically significant for GD. This may suggest that the disruption associated with having a letter mask preview of target word 2 did not affect the processing at Target Word 1 itself, but instead affected the parafoveal processing of Target Word 2 at the point of fixating Target Word 1. Again, similar to the results on Pre-Target, I further conducted Bayes Factor calculations in the same manner for FFD, SFD, GD, RP, GPT and ROut specifically to examine the effect of having target word 2 masked in preview for all six measures as well as FFD, SFD, RP and ROut only for the interaction between the two factors (see Table 4.3). Again, in the same way as for Pre-Target, there was no clear evidence in favour of the alternative hypothesis ($BF \leq 0.52$).

The observed effects on Target Word 1 are largely consistent with the findings by Angele et al. (2008) with one exception. Having a letter mask preview for target word 2 did influence GD, GPT and TVT on Target Word 1 only when readers received an identical preview for target word 1 in the present experiment while it had no influence on any measures in the study by Angele et al. (2008). Importantly, despite this difference, the results from both studies seem to provide evidence that parafoveal processing must at least to an extent be serial. This is because, according to both E-Z Reader (Reichle et al., 1998; Reichle, 2011) and Über Reader (Reichle, 2021), words are processed serially such that once a word is identified, attention is shifted to the next word readers aim to process. As discussed before (see Schotter et al., 2014) it is possible that both word *n* and word *n+1* are identified rapidly enough that parafoveal processing word *n+2* can begin at the point of fixating word *n*. The lack of any parafoveal-on-foveal and parafoveal preview n+2 effects observed by Angele et al. can thus be explained such that at the point of fixating the Pre-Target, readers were not able to parafoveally identify the first target word quickly enough for parafoveal processing of the preview of the second target word to commence before the eyes crossed the boundary. In the present study, I observed disruption to reading on Target Word 1 associated with having a masked preview of target word 2 only when the first target word was not masked in preview. Hence, the present results suggest that readers were only able to start parafoveally processing target word 2 at the point of fixating the Pre-Target only when lexical identification of target word 1 was possible (i.e., its preview was not masked).

This is problematic for the parallel processing accounts of SWIFT (Engbert et al., 2002, 2005), SEAM (Rabe et al., 2024) and OB1 Reader (Snell, van Leipsig, et al., 2018) since according to these models, attention is distributed to multiple words simultaneously during a given fixation. Consequently, the second target word should have received some parafoveal processing at the point of fixating the Pre-Target regardless of whether target word 1 was masked or not in preview. Conversely, this pattern of effects may fit within the serial framework of E-Z Reader (Reichle, 2011) and Über Reader (Reichle, 2021) since it seems that readers could only parafoveally process target word 2 when they were also able to effectively parafoveally process target word 1 first.

On Target Word 2, the effect of target word 1 being masked was not significant for any measure (see Figure 4.4). This pattern indicates that by the time readers fixated on Target Word 2, they had largely recovered from the disruption they experienced upon fixating Target Word 1 in the event that the first target word was masked in preview. The effect of target word 2 mask was only significant for ROut (and no other measures) such that readers were slightly more likely to make a regressive saccade out of Target Word 2 when it had been masked in preview versus when it was not masked in preview. I note that for all conditions, the probability to regress out of Target Word 2 was quite low ($M \leq 12\%$) and the effect seems to be relatively small in size (approximately 3%). Hence, I tentatively propose that this effect may be spurious and not necessarily reflect any disruption to processing associated with having a masked preview of the second target word. At the very least, this specific significant effect should be interpreted with caution. Lastly, there was a significant interaction on TVT only such that the effect of target word 2 mask was slightly larger when target word 1 was not masked versus when it was masked in preview. Given the lack of any earlier effects on Target Word 2, this result is not straightforward to interpret. It may be that the interaction on TVT, similar to the effect of target word 2 mask on ROut, is spurious. Assuming this to be the case, these results seem to indicate that readers experienced little to no difficulty processing Target Word 2 across all four conditions. Hence disruption to processing associated with having one or both target words masked in preview seemed to affect processing reliably only at the Target Word 1 region. Overall, the results from the analyses for the three regions of interest are predominantly consistent with serial over parallel processing assumptions.

4.4.5. Single-factor models

On the Pre-Target region, there was a significant effect of the two target words being transposed in preview on TVT and RIn only with no other significant effects on any other measures (see Figure 4.2). As with the interactive models, I calculated Bayes Factors in the same manner to assess to what extent the current data support the notion that reading times and fixation probabilities on the Pre-Target prior to crossing the invisible boundary were comparable across the 3 types of sentences investigated with this set of models (see Table

4.4). In the same fashion as with the interactive models, the Bayes Factors (*BF* ≤ 0.42) provided further support that there were no significant parafoveal-on-foveal effects on processing at the Pre-Target. Readers made more regressions into the Pre-Target and consequently, TVT was longer when the two target words were transposed versus when they were not transposed in preview. Similar to the results for the Pre-Target region from the interactive models, this seems to suggest that readers experienced processing difficulty after their eyes had crossed the boundary when the order of the two target words changed from being transposed in preview to being correct in foveal vision. Furthermore, there was no difference in TVT and RIn between when the two target words were transposed versus masked in preview suggesting that disruption to processing by a transposed-word preview or a masked preview in comparison to the identity preview condition was comparable. The lack of early parafoveal-on-foveal effects at the Pre-Target region for the transposed target word previews and the masked target word previews is inconsistent with the SWIFT (Engbert et al., 2002, 2005) and SEAM (Rabe et al., 2024) parallel processing accounts, since both models would predict that readers should parafoveally process both target word 1 and target word 2 at the point of fixating the Pre-Target region meaning that they should detect the presence of letter masks for both target words or the ungrammaticality created by the transposition of the two target words parafoveally. Clearly, this did not happen. In the case of the transposed target word previews, the transposition rendered the sentence ungrammatical at the Target Word 1 region in the sentence meaning that at the very least, readers should have detected that the upcoming word in the parafovea was ungrammatical even if they were not fully aware that the two target words were parafoveally transposed.

The lack of any difference in the disruption to processing between the transposed targets and the masked targets previews is problematic for OB1 Reader (Snell, van Leipsig, et al., 2018). According to this model, the detection of a word transposition should not affect local eye movement measures on the transposed words or the word preceding the transposition. What is more, according to the model, readers integrate orthographic information across words, meaning that the masked target preview condition should cause significant disruption (certainly more than the transposition masks) because the letter string masks for both target words had no orthographic correspondence with the target words themselves. Therefore, it stands to reason, that disruption associated with the transposed target previews should be reduced in comparison to the disruption associated with the masked target previews (see Section 1.6.3). Finally, these results are consistent with the serial processing accounts of E-Z Reader (Reichle, 2011) and Über Reader (Reichle, 2021) since readers experienced comparable disruption to processing associated with either type of preview change only after the previews changed to the target words (i.e., after the eyes crossed the boundary).

On Target Word 1, the effect of the transposed target words preview remained significant for all measures (see Figure 4.3) aside from RIn, for which it approached significance. Reading times as well as the probability to refixate and to regress out of the Target Word 1 region were inflated when the two target words were transposed versus when they were identical in preview. This is consistent with a preview change cost and indicates that readers detected the preview change quite rapidly, and disruption was relatively long-lasting. Additionally, disruption to processing was comparable across the two types of preview change for FFD, SFD, GD, RP, ROut and TVT as that the two types of preview change created comparable disruption to processing on Target Word 1 for all these measures.

At least for first-pass and early reading (before the eyes have moved to the right in the sentence) measures, the transposed-word preview and the masked target word previews were comparable (see Figure 4.3). These results were further confirmed by Bayes Factor calculations (see Table 4.4) for FFD, SFD, GD, RP and ROut, ($BF \leq 0.26$) There was some evidence for a difference between the transposed target word preview and the masked target word preview changes at least with respect to rereading behaviour. Firstly, on GPT, there was a significant difference between the transposed target word preview condition and the masked target word preview condition such that masked target word previews resulted in longer GPT than did the transposed target word previews. This is consistent with the findings by Rayner, Angele, et al. (2013) who showed that GPT on the first word following the boundary is inflated more when the two post-boundary words are masked by unrelated words in preview versus when they are transposed and are grammatical given previous context. The authors interpreted this as evidence that a transposed target word preview has more orthographic overlap with the target words themselves than a preview consisting of two unrelated valid words. In the present experiment, the transposed target word previews meant that the two target words were not only presented out of order parafoveally, but also that they were ungrammatical in parafoveal vision. However, it is also true that the transposed target word previews had more overlap with the target words themselves than was the case for the masked target previews (recall, that there was no orthographic overlap in this case). Therefore, it seems reasonable to interpret the pattern observed on GPT as evidence that disruption to processing increases as the orthographic overlap between preview and targets decreases.

Finally, for the Target Word 1 region, the effect of having masked compared to transposed target words in preview for RIn approached significance. This effect was, however, in the opposite direction to the one observed for GPT. Readers were less likely to make a regressive saccade into Target Word 1 when the two targets were masked than when they were transposed in preview. I tentatively suggest that this pattern may indicate that the transposed target word preview caused disruption to processing that lasted longer than the disruption to processing caused by the masked target word preview. This could potentially be explained by the fact that in the transposed target word preview condition, readers received a preview of target word 2 as the upcoming word in the parafovea twice, once before their eyes crossed the boundary, and a second time when they fixated Target Word 1. Hence, while processing on Target Word 1 may have been disrupted by the preview change, processing of Target Word 2 could be enhanced. Consequently, readers might need less opportunity to foveally process target word 2 and seek to obtain more processing time on target word 1, leading to inflated regression rates into Target Word 1. Conversely, the masked target word previews would provide no such benefit to parafoveal processing of the Target Word 2 region since the second target word was masked by a string of orthographically unrelated letters in preview and remained in its correct position. I do, though, note that the trends observed on RIn for both the transposed target word reviews and the masked target word previews failed to reach statistical significance, and therefore, may not indicate meaningful differences in processing costs associated with transposing versus masking the two target words with strings of letters in preview.

This set of results is consistent with the serial accounts of E-Z Reader (Reichle, 2011) and Über Reader (Reichle, 2021) as well as the parallel gradient accounts of SWIFT (Engbert et al., 2002, 2005) and SEAM (Rabe et al., 2024) since there was a clear preview change cost to processing at the Target Word 1 region for both types of preview change. With respect to OB1 Reader (Snell, van Leipsig, et al., 2018), however, these results are problematic. This is because the model would predict that a masked target word preview change should produce significantly more disruption to processing than a transposed target word preview change, for the same reasons as discussed for the results observed at the pre-target region. There was some evidence in favour of such a prediction on GPT. However the lack of any difference between the transposed Targets and masked Targets previews for FFD, SFD, GD, RP, ROut and TVT, as well as the patterns observed for RIn, suggest that the pattern observed on GPT is likely produced by the increased overlap between the target words in parafoveal versus foveal vision in the transposed preview relative to the masked preview conditions, rather than because of the mechanisms proposed by OB1 Reader.

On Target Word 2, the effect of having a transposed target word preview was only significant for GPT, ROut and TVT, such that readers made more regressive saccades out of the Target Word 2 region, and had longer GPT, and TVT in the transposed target versus the identity preview condition. In addition, GPT, ROut and TVT were significantly lower in the masked target word preview condition than in the transposed target word preview condition. Furthermore, there was no differential processing disruption on FFD, SFD, GD, RP and RIn for either of the two preview change conditions.

These three patterns of effects clearly suggest that the disruption to processing associated with the transposition of the two target words in preview was prolonged compared to the disruption caused by masking the two target words with letter strings in preview. Furthermore, I propose that these effects fit with the pattern observed for RIn at Target Word 1. Therefore, the inflated GPT, ROut and TVT on Target Word 2 may have been due to the difficulty readers experienced with processing Target Word 1, rather than any difficulty in processing Target Word 2. That is to say, parafoveal processing of Target Word 2 was potentially enhanced when the two target words were transposed in comparison to masked by letter strings in preview. Consequently, ROut, GPT, and TVT for Target Word 2 as well as RIn for Target Word 1 were inflated for transposed target words compared to the identity and masked target word preview conditions. This set of findings, similarly to the findings on Pre-Target and Target Word 1 offers support for the notion that words are processed serially and sequentially rather than in parallel.

4.4.6. Results summary

The results from both the interactive models and the single-factor models indicate that sensitivity to the characteristics of the Target Word 2 region in parafoveal vision at the point of fixating the Pre-Target region was limited. The disruption to processing on Target Word 1 when target word 2 was masked, and target word 1 was not masked in preview, suggests, at least to some degree, that parafoveal processing is serial. In addition, only processing of Target Word 1 after the eyes had crossed the boundary was reliably disrupted by all four types of preview change used in the present experiment. Furthermore, the observed patterns on Target Word 1 when only the first target word was masked, when both targets were masked and when the two targets were transposed in preview were all comparable (with the exception of GPT). Consequently, the primary source of the disruption observed in the present experiment was likely the invalid preview of target word 1 across those three (non-identical preview) experimental conditions.

Table 4.2.

Descriptive statistics for all measures

Measure/ Condition	Identity	Target Word 1 Masked Target Word 2 Not Masked	Target Word 1 Not Masked Target Word 2 Masked	Masked Targets	Transposed Targets
			Pre-Target		
FFD	227 (69)	231 (69)	229 (70)	232 (71)	231 (69)
SFD	228 (65)	229 (65)	231 (69)	230 (64)	230 (69)
GD	262 (106)	265 (107)	266 (108)	270 (100)	266 (103)
RP	19 (39)	18 (38)	19 (39)	21 (41)	20 (40)
GPT	291 (138)	297 (145)	286 (132)	297 (135)	295 (144)
ROut	10 (30)	10 (29)	8 (27)	9 (29)	8 (27)
TVT	318 (180)	351 (192)	319 (166)	348 (183)	342 (189)
RIn	11 (31)	24 (42)	12 (32)	22 (42)	19 (39)
			Target Word 1		
FFD	237 (72)	269 (80)	246 (78)	268 (86)	257 (88)
SFD	238 (67)	279 (75)	255 (80)	281 (82)	270 (88)
GD	265 (100)	314 (107)	281 (105)	310 (104)	302 (120)
RP	15 (36)	23 (42)	18 (39)	22 (41)	21 (41)
GPT	282 (128)	379 (196)	311 (154)	366 (182)	348 (183)
ROut	5 (23)	18 (39)	7 (26)	17 (38)	14 (34)
TVT	311 (151)	374 (166)	333 (157)	370 (166)	375 (184)
RIn	12 (32)	10 (30)	12 (33)	12 (33)	15 (36)
			Target Word 2		
FFD	231 (70)	234 (76)	232 (71)	234 (73)	239 (79)
SFD	238 (74)	236 (76)	233 (70)	233 (67)	243 (80)
GD	261 (102)	263 (105)	261 (99)	257 (96)	267 (107)
RP	9 (28)	9 (29)	11 (32)	12 (32)	17 (38)
GPT	289 (140)	294 (148)	297 (154)	295 (148)	323 (179)
ROut	9 (28)	9 (29)	12 (32)	12 (32)	17 (38)
TVT	304 (145)	317 (159)	315 (157)	310 (154)	329 (164)
RIn	11 (32)	12 (32)	12 (32)	9 (29)	11 (31)

Note. Values for FFD, SFD, GD, GPT and TVT are given in milliseconds while values for RP, ROut and RIn are given in percentages.

Table 4.3.

Fixed effects estimates from the generalised mixed effects interactive models.

Measure/ Condition	Intercept				Target Word 1 Mask (Masked versus Not Masked)					Target Word 2 Mask (Masked versus Not Masked)					Interaction				
Pre-Target																			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	234.81	4.28	**54.91**	[226.43,243.19]	1.92	3.14	0.61	0.02	[-4.24,8.08]	1.42	3.16	0.45	0.02	[-4.78,7.62]	-6.45	5.13	-1.26	0.05	[-16.50,3.61]
SFD	244.77	6.67	**36.72**	[231.71,257.84]	-2.97	4.01	-0.74	0.02	[-10.82,4.89]	2.17	3.97	0.55	0.02	[-5.62,9.96]	-5.77	6.95	-0.83	0.04	[-19.38,7.84]
GD	272.86	7.37	**37.01**	[258.41,287.31]	1.05	4.47	0.24	0.02	[-7.71,9.81]	5.20	4.63	1.12	0.03	[-3.89,14.28]	-3.62	3.61	-0.55	0.03	[-16.57,9.33]
RP	-1.72	0.13	**-13.45**	[-1.97, -1.47]	0.01	0.08	0.17	0.05	[-0.15,0.18]	0.14	0.08	1.68	0.20	[-0.02,0.30]	0.28	0.17	1.66	0.38	[-0.05,0.60]
GPT	303.04	5.70	**53.20**	[291.88,314.21]	4.50	5.01	0.90	0.03	[-5.32,14.32]	-1.65	4.57	-0.36	0.02	[-10.62,7.31]	-3.37	7.18	-0.47	0.03	[-17.45,10.71]
ROut	-2.54	0.12	**-22.10**	[-2.77, -2.32]	-0.02	0.14	-0.13	0.05	[-0.29,0.25]	-0.16	0.11	-1.44	0.12	[-0.37,0.06]	0.20	0.22	0.92	0.13	[-0.23,0.63]
TVT	344.74	5.99	**57.60**	[333.01,356.47]	29.19	4.85	**6.02**	N/A	[19.69,38.69]	3.39	4.85	0.70	N/A	[-6.13,12.91]	-8.61	6.89	-1.25	N/A	[-22.11,4.88]
RIn	-1.96	0.14	**-14.41**	[-2.23, -1.69]	0.92	0.13	**7.08**	N/A	[0.67,1.17]	0.03	0.13	0.29	N/A	[-0.16,0.21]	-0.26	0.19	-1.39	N/A	[-0.62,0.11]
Target Word 1																			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	258.49	5.34	**48.44**	[248.03,268.95]	28.49	4.77	**5.98**	N/A	[19.15,37.83]	2.74	3.52	0.78	0.02	[-4.16,9.63]	-9.98	6.16	-1.62	0.09	[-22.06,2.09]
SFD	276.74	7.75	**35.73**	[261.56,291.92]	37.37	6.52	**5.73**	N/A	[24.60,50.14]	9.11	5.18	1.76	0.09	[-1.04,19.25]	-11.56	9.83	-1.18	0.07	[-30.83,7.70]
GD	296.46	6.08	**48.74**	[284.54,308.38]	42.00	5.64	**7.45**	N/A	[30.95,53.05]	5.58	4.38	1.27	0.03	[-3.01,14.16]	-18.47	5.45	**-3.39**	N/A	[-29.15, -7.78]
RP	-1.60	0.10	**-15.85**	[-1.80, -1.40]	0.48	0.13	**3.88**	N/A	[0.24,0.73]	0.12	0.08	1.43	0.15	[-0.05,0.29]	-0.26	0.17	-1.56	0.35	[-0.59,0.07]
GPT	344.72	4.99	**69.11**	[334.94,354.50]	86.75	6.62	**13.11**	N/A	[73.79,99.72]	9.14	4.45	2.04	0.10	[0.34,17.93]	-46.01	5.29	**-8.70**	N/A	[-56.37, -35.65]
ROut	-2.45	0.14	**-17.94**	[-2.71, -2.18]	1.11	0.15	**7.25**	N/A	[0.81,1.41]	0.10	0.12	0.88	0.07	[-0.12,0.33]	-0.43	0.23	-1.86	0.52	[-0.88,0.02]
TVT	361.01	9.87	**36.60**	[341.68,380.35]	56.46	9.10	**6.21**	N/A	[38.63,74.28]	9.05	6.56	1.38	N/A	[-3.81,21.91]	-26.03	10.22	**-2.55**	N/A	[-46.06, -6.01]
RIn	-2.20	0.10	**-21.62**	[-2.40, -2.00]	-0.10	0.10	-1.01	N/A	[-0.30,0.10]	0.12	0.10	1.21	N/A	[-0.08,0.32]	0.16	0.20	0.80	N/A	[-0.24,0.56]
Target Word 2																			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	234.73	4.92	**47.69**	[225.08,244.38]	1.79	3.02	0.59	N/A	[-4.13,7.72]	0.18	2.73	0.06	N/A	[-5.17,5.52]	-1.76	4.55	-0.39	N/A	[-10.68,7.17]
SFD	241.68	6.23	**38.82**	[229.48,253.88]	0.70	3.46	0.20	N/A	[-6.09,7.49]	-5.25	3.34	-1.57	N/A	[-11.79,1.29]	0.75	5.22	0.14	N/A	[-9.49,10.98]
GD	262.28	5.51	**47.61**	[251.48,273.08]	0.98	3.39	0.29	N/A	[-5.67,5.23]	-3.44	4.43	-0.78	N/A	[-12.12,5.23]	-4.36	4.80	-0.91	N/A	[-13.77,5.06]
RP	-2.06	0.12	**-17.43**	[-2.29, -1.83]	-0.12	0.12	-1.01	N/A	[-0.35,0.11]	-0.10	0.10	-0.98	N/A	[-0.29,0.10]	-0.10	0.20	-0.53	N/A	[-0.48,0.28]
GPT	298.71	6.64	**45.02**	[285.71,311.72]	2.26	5.96	0.38	N/A	[-9.42,13.93]	4.68	6.20	0.76	N/A	[-7.47,16.83]	-8.79	10.35	-0.85	N/A	[-29.08,11.51]
ROut	-2.39	0.11	**-21.60**	[-2.60, -2.17]	0.04	0.11	0.36	N/A	[-0.18,0.26]	0.32	0.11	**2.85**	N/A	[0.10,0.54]	-0.05	0.22	-0.24	N/A	[-0.49,0.38]
TVT	318.19	7.37	**43.19**	[303.75,332.63]	3.71	5.20	0.71	N/A	[-6.49,13.91]	3.78	4.89	0.77	N/A	[-5.81,13.38]	-21.02	6.31	**-3.33**	N/A	[-33.40, -8.64]
RIn	-2.36	0.11	**-20.82**	[-2.59, -2.14]	-0.09	0.13	-0.72	N/A	[-0.34,0.16]	-0.08	0.11	-0.72	N/A	[-0.29,0.13]	-0.26	0.22	-1.19	N/A	[-0.68,0.17]

Note. Significant terms are presented in bold, and terms approaching significance are underlined. All *p* values were adjusted via the Bonferroni correction with the summary(glht(<modelName>), test = adjusted(“bonferoni”)) function. *BF* refers to the Bayes factor calculated via an R adaptation of the Dienes calculator (Dienes, 2008) by Silvey et al. (2024).

^Confidence intervals (Conf Int) are calculated with the *confint* function in R with method = *Wald* at the 2.5 percentile and 97.5 percentile

Table 4.4.

Fixed effects estimates from the generalised mixed effects single-factor models.

Measure/ Condition	Intercept				Preview Type (Transposed Targets versus Identity)					Preview Type (Masked Targets versus Transposed Targets)				
Pre-Target														
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	233.36	5.64	**41.41**	[222.31,244.40]	5.40	3.57	1.51	0.05	[-1.59,12.39]	-1.93	3.19	-0.61	0.02	[-8.18,4.31]
SFD	241.28	5.97	**40.42**	[229.59,252.98]	-0.27	5.45	-0.05	0.02	[-10.96,10.41]	-0.05	5.37	-0.01	0.02	[-10.58,10.48]
GD	270.23	6.51	**41.52**	[257.47,282.98]	5.52	3.81	1.45	0.04	[-1.95,12.98]	2.65	4.82	0.55	0.02	[-6.81,12.10]
RP	-1.63	0.13	**-12.67**	[-1.88, -1.37]	0.04	0.12	0.35	0.08	[-0.19,0.27]	0.11	0.12	0.98	0.11	[-0.11,0.34]
GPT	301.56	7.32	**41.23**	[287.23,315.90]	4.42	5.41	0.82	0.03	[-6.19,15.03]	1.54	7.76	0.20	0.03	[-13.66,16.74]
ROut	-2.64	0.14	**-18.37**	[-2.92, -2.36]	-0.31	0.16	-1.97	0.42	[-0.63, -0.00]	0.23	0.16	1.41	0.17	[-0.09,0.54]
TVT	344.58	7.52	**45.85**	[329.85,359.31]	23.11	4.74	**4.88**	N/A	[13.82,32.40]	5.31	4.74	1.12	N/A	[-3.98,14.60]
RIn	-1.94	0.14	**-14.21**	[-2.21, -1.67]	0.82	0.17	**4.94**	N/A	[0.50,1.15]	0.15	0.15	1.03	N/A	[-0.14,0.44]
Target Word 1														
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	258.67	5.95	**43.45**	[247.01,270.34]	21.38	4.82	**4.44**	N/A	[11.93,30.82]	8.72	5.19	1.68	0.08	[-1.46,18.90]
SFD	278.29	7.78	**35.77**	[263.04,293.53]	30.70	7.75	**3.96**	N/A	[15.51,45.89]	12.66	7.38	1.71	0.12	[-1.81,27.13]
GD	296.60	8.07	**36.77**	[280.79,312.41]	38.29	6.59	**5.81**	N/A	[25.38,51.20]	5.73	6.51	0.88	0.03	[-7.04,18.49]
RP	-1.62	0.11	**-14.96**	[-1.83, -1.41]	0.43	0.13	**3.42**	N/A	[-0.18,0.68]	0.10	0.12	0.82	0.10	[-0.13,0.32]
GPT	341.09	9.34	**36.51**	[322.78,359.40]	72.86	11.07	**6.58**	N/A	[51.16,94.55]	20.80	8.17	**2.55**	N/A	[4.80,36.80]
ROut	-2.49	0.15	**-17.08**	[-2.78, -2.21]	1.10	0.25	**4.33**	N/A	[0.60,1.60]	0.27	0.16	1.68	0.26	[-0.04,0.58]
TVT	363.28	9.75	**37.26**	[344.17,382.39]	63.62	7.77	**8.19**	N/A	[48.40,78.84]	1.27	9.94	0.13	N/A	[-18.22,20.76]
RIn	-2.00	0.09	**-21.66**	[-2.18, -1.82]	0.31	0.14	2.26	N/A	[0.04,0.58]	-0.30	0.13	-2.20	N/A	[-0.56, -0.03]
Target Word 2														
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*	*β*	*SE*	*t(z)-value*	*BF*	*Conf Int*
FFD	236.58	4.64	**51.00**	[227.49,245.68]	8.34	4.08	2.04	N/A	[0.33,16.34]	-7.08	4.30	-1.65	N/A	[-15.51,1.35]
SFD	242.88	6.14	**39.53**	[230.84,254.92]	7.59	5.38	1.41	N/A	[-2.95,18.13]	-10.40	5.51	-1.89	N/A	[-21.19,0.39]
GD	263.53	6.88	**38.30**	[250.04,277.01]	5.78	4.57	1.26	N/A	[-3.19,14.74]	-9.82	5.18	-1.90	N/A	[-19.98,0.33]
RP	-2.05	0.12	**-16.63**	[-2.29, -1.81]	0.01	0.14	0.05	N/A	[-0.26,0.27]	-0.21	0.14	-1.51	N/A	[-0.48,0.06]
GPT	308.12	8.33	**37.00**	[291.80,324.44]	40.09	6.96	**5.76**	N/A	[26.45,53.72]	-30.91	7.42	**-4.17**	N/A	[-45.45, -16.37]
ROut	-2.15	0.11	**-20.23**	[-2.36, -1.95]	0.86	0.17	**5.03**	N/A	[0.53,1.20]	-0.46	0.16	**-2.82**	N/A	[-0.77, -0.14]
TVT	319.90	8.89	**36.00**	[302.48,337.32]	22.28	5.65	**3.95**	N/A	[11.22,33.35]	-17.40	5.59	**-3.12**	N/A	[-28.35, -6.46]
RIn	-2.38	0.13	**-18.70**	[-2.63, -2.13]	-0.08	0.15	-0.49	N/A	[-0.37,0.22]	-0.11	0.16	-0.71	N/A	[-0.42,0.20]

Note. Significant terms are presented in bold, and terms approaching significance are underlined. All *p* values were adjusted via the Bonferroni correction with the summary(glht(<modelName>), test = adjusted("bonferoni")) function. *BF* refers to the Bayes factor calculated via an R adaptation of the Dienes calculator (Dienes, 2008) by Silvey et al. (2024).

^Confidence intervals (Conf Int) are calculated with the *confint* function in R with method = *Wald* at the 2.5 percentile and 97.5 percentile

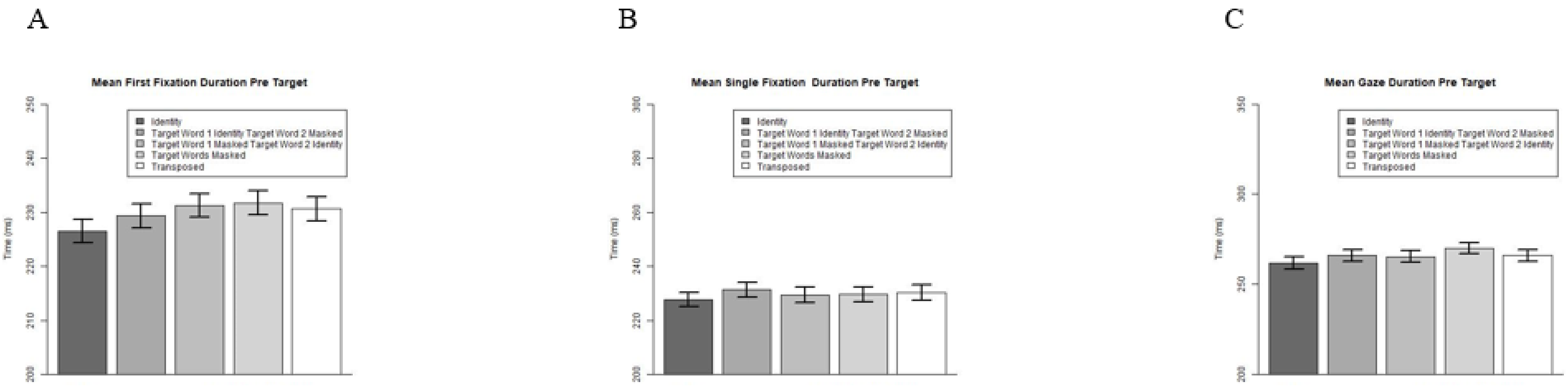


Figure 4.2. Mean first-pass reading times on Pre-Target

Panel A shows the mean first fixation duration on the Pre-Target region; Panel B shows the mean single fixation duration on the Pre-Target region; Panel C shows the mean gaze duration on the Pre-Target region.

Error bars represent standard errors.

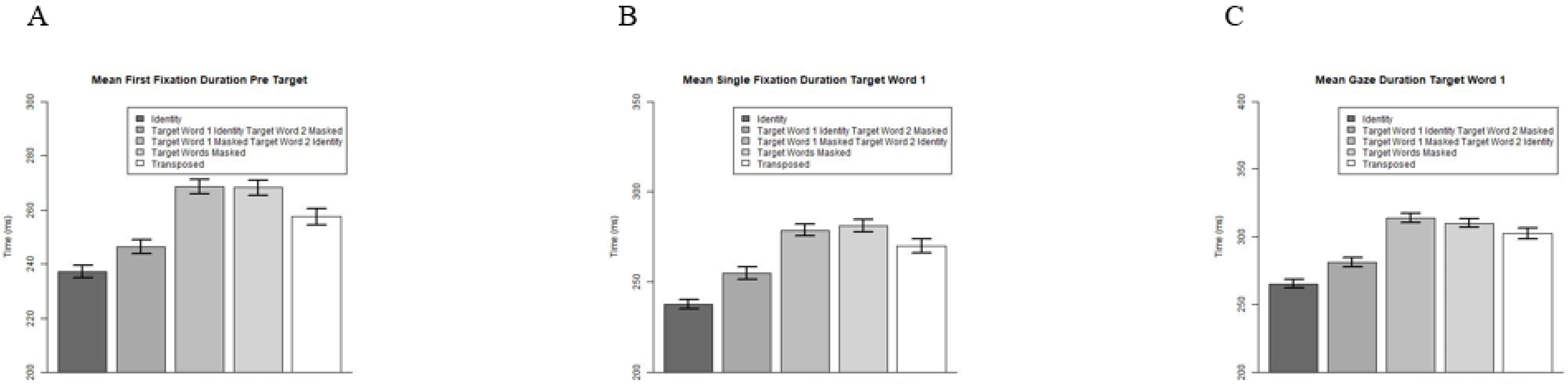


Figure 4.3. Mean first-pass reading times on Target Word 1

Panel A shows the mean first fixation duration on the Target Word 1 region; Panel B shows the mean single fixation duration on the Target Word 1 region; Panel C shows the mean gaze duration on the Target Word 1 region.

Error bars represent standard errors.

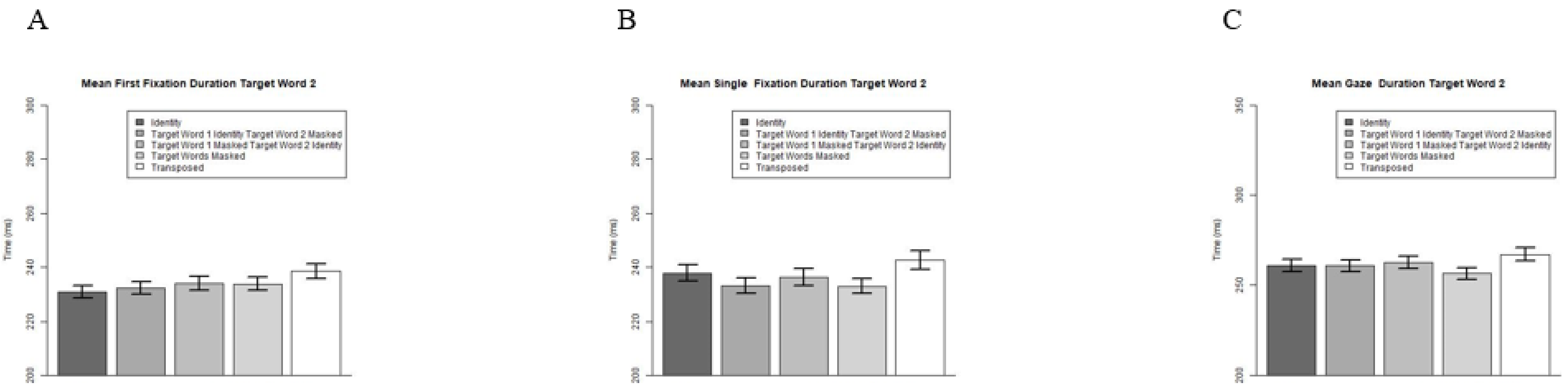


Figure 4.4. Mean first-pass reading times on Target Word 2

Panel A shows the mean first fixation duration on the Target Word 2 region; Panel B shows the mean single fixation duration on the Target Word 2 region; Panel C shows the mean gaze duration on the Target Word 2 region.

Error bars represent standard errors.

4.5. Discussion

The present experiment was aimed at investigating whether readers are sensitive to the characteristics of word *n+2* at the point of fixating word *n* via the boundary paradigm (Rayner, 1975). Further, in the event that there was indeed some parafoveal sensitivity to word *n+2* did that sensitivity extend to the orthographic and/or (post) lexical characteristics of word *n+2*. I manipulated the preview of one or two target words such that one or both targets could be masked by letter strings consisting entirely of consonants that could not form a valid word, or they could be transposed resulting in an ungrammaticality. I ensured that the two target words and the pre-target word were controlled for length and frequency. Furthermore the two target words were selected to be equally (un)predictable, in order to ensure that any differences on the pre-target or target words were driven by the experimental manipulation. My first aim was to investigate the existence of parafoveal-on-foveal and parafoveal preview n+2 effects due to having letter masked previews of one or both target words (e.g., Angele et al., 2017; see Vasilev & Angele, 2017). Moreover, I aimed to explore differences in parafoveal sensitivity to word transpositions in comparison to letter masks of word *n+1* and word *n+2* in preview.

The question of n+2 parafoveal sensitivity has been one of the primary ways to test if words are lexically processed in parallel as per SWIFT (Engber678 et al., 2002, 2005) and SEAM (Rabe et al., 2024) or OB1 Reader (Snell, van Leipsig, et al., 2018), or if they are processed serially in line with E-Z Reader (Reichle et al., 1998; Reichle, 2011) and Über Reader (Reichle, 2021). In particular, if processing proceeds as per SWIFT and SEAM, I expected to find effects of the preview of target word 2 as early as at the Pre-Target region and independently from any effects associated with the preview of target word 1. Furthermore, parafoveal word transpositions should also cause disruption to processing starting at the Pre-Target. Conversely, according to OB1 Reader, having a letter mask for one or both target words should cause disruption on the Pre-Target region while having a transposition in preview should have no effect on the Pre-Target. On Target Word 1 and/or Target Word 2, having a transposition in preview should cause significantly less disruption to processing than having letter masks for one or both targets in preview. Finally, based on E-Z Reader and Über Reader, there should be no parafoveal-on-foveal effects of any type of preview change on the Pre-Target. Further, there should be no disruption to processing on the first target word associated with the preview of the second target word (see Angele et al., 2008). Additionally, having a transposed word pair in preview should be disruptive to processing, potentially even more so than having letter masks for both target words in preview. As discussed above, this is because readers may be able to parafoveally process the upcoming word in the parafovea to a greater extent when the two targets are transposed versus when they are masked.

Given these predictions, the results lead to several important conclusions. Firstly, there were no parafoveal-on-foveal effects on the Pre-Target region for any type of preview change which was further confirmed via Bayes Factor calculations for early reading time measures (see Section 4.4.4 & 4.4.5). This indicates that parafoveal sensitivity to target word 2 at the point of fixating the Pre-Target is limited and is potentially problematic for all three parallel accounts (although see above for an in-depth discussion). Secondly, there was significant disruption to processing on Target Word 1 which was, for most measures (aside from GPT and potentially RIn) analysed, comparable between having both target words masked versus having the two targets transposed in preview. In addition, disruption was comparable when readers were provided with a masked preview for only target word 1 versus a masked preview for both target words. Consequently, it is clear that the primary source of disruption on Target Word 1 was associated with having an invalid preview of the first target regardless of whether the preview of target word 2 was valid or not. This is, overall, consistent only with the serial processing accounts of E-Z Reader and Über Reader and is potentially problematic for all three parallel accounts discussed so far (SWIFT; SEAM; OB1 Reader).

The third key finding is that the parafoveal preview of the second target word did influence processing on Target Word 1. However that influence was delayed and smaller in comparison to the influence of the preview of target word 1 itself. Furthermore, having a masked preview of target word 2 only seemed to disrupt processing on Target Word 1 when the first target itself was not masked in preview. This potentially suggests that readers were able to parafoveally process the second target word at the point of fixating the Pre-Target only when they were also able to parafoveally process (and potentially identify) target word 1 as well. Hence, I would argue that this finding can be taken as evidence that parafoveal processing is likely, at least to an extent, serial. Therefore, I posit that this finding fits more within the serial processing frameworks of E-Z Reader and Über Reader (see Schotter et a., 2014; see Section 4.4.6) over any of the three parallel models discussed in this chapter. This is because, according to all three parallel models, any effects associated with the preview of target word 2 should be independent from the effects associated with the preview of target word 1.

The final key point is that by the time readers moved their eyes onto Target Word 2, they had largely recovered from any disruption they had experienced on Target Word 1 associated with any of the four types of preview change (see Section 4.4.6). Again, similar to the patterns observed on the Pre-Target, this pattern is potentially problematic for the three parallel models (SWIFT; SEAM; OB1 Reader). In contrast, this finding fits entirely with the predictions of the E-Z Reader and Über Reader models.

4.5.1. Limitations of the present study

There are two important considerations to make with respect of the current results. Firstly, as discussed before (see Section 1.5.2; Vasilev & Angele, 2017), n+*2* parafoveal preview effects are very small. Hence, it is possible that the current study was not sufficiently statistically powered to examine such small effects. This is partly because the study was initially designed as a co-registered eye-tracking and electroencephalography experiment. Consequently, I did not conduct an a priori power analysis to estimate the sample size required to obtain an n+*2* parafoveal preview or parafoveal-on-foveal effect. In order to mitigate for this issue, I have conducted Bayes Factors calculations (see Section 4.4.4 & 4.4.5). In all cases, these calculations confirmed the non-significant effects. Despite this, it may be beneficial for future studies aiming to investigate n+2 parafoveal preview and parafoveal-on-foveal effects to aim for larger samples by increasing the number of observations per participant, and/or increasing the number of participants.

Secondly, similar to Experiment 2 (see Section 3.5.1), in the present study, Target Word 2 was always, at least partially, in peripheral vision at the point of fixating on the Pre-Target region. This is important because visual acuity in the periphery is significantly lower than visual acuity in the parafovea and peripheral vision is subject to stronger crowding effects than parafoveal vision (see Section 1.4.1; Rosenholtz, 2016). This, again, was inevitable given the characteristics of the materials in the present study (see Section 4.3.4). This, however, is not only an issue of the present study, but also of most other studies of n+*2* effects in alphabetic spaced languages (see Vasilev & Angele, 2017). Consequently, it may be beneficial for future investigations of n+*2* effects to consider more directly, to what extent word *n+2* is actually in parafoveal versus peripheral vision at the point of fixating the pre-boundary word.

4.6. Conclusion

In summary, the present eye movement results are relevant to the parallel versus serial processing debate in the eye-tracking reading research (see Zang, 2019 for a discussion). The effects observed across the three regions of interest in both sets of analyses I conducted do clearly support the notion that words are (both foveally and parafoveally) processed serially as per E-Z Reader and Über Reader and are overall consistent with the conclusions of Vasilev and Angele (2017). Furthermore, it is unlikely that processing occurs as per SWIFT, SEAM or OB1 Reader. This is because any effects associated with the parafoveal preview of Target Word 2 were only visible when the parafoveal preview of Target Word 1 was left intact and were substantially smaller than the effects associated with changes to the preview of Target Word 1 itself.

Chapter 5. General Discussion

The present thesis aimed to examine the recently established *Transposed-Word* effect (Mirault et al., 2018) and how word transpositions may affect reading via three eye-tracking reading experiments. This topic was chosen because of how relevant *Transposed-Word* effects have become to the serial versus parallel lexical processing debate primarily in relation to the OB1 Reader model (Snell, van Leipsig, et al., 2018; see Section 1.6.3). Across all three experiments, the two target words which were sometimes transposed (in preview and/or foveal vision) with the immediately preceding pre-target word were carefully selected in accordance with both a strict interpretation of the principles of the OB1 Reader, and existing eye movement reading research (see Clifton et al., 2016). This was done in order to maximise the likelihood that participants failed to detect the word transposition and processed the target words as would be predicted by OB1 Reader (i.e., readers would fail to detect the transposition). Therefore, the three experiments were optimally designed to test the predictions of the OB1 Reader model.

Each of the three experiments of the current thesis has been discussed in depth in its respective chapter (Experiment 1 – Chapter 2; Experiment 2 – Chapter 3; Experiment 3 – Chapter 4). Therefore, in this chapter I will not summarise any of the three studies. Instead, in the subsequent section I will outline and discuss the key conclusions that can be drawn from the three experiments that comprised my Ph.D. project. I will further consider the implications of each conclusion for the five computational models of oculomotor control during reading discussed in this thesis, namely E-Z Reader (Reichle et al., 1998; Reichle, 2011), Über Reader (Reichle, 2021); SWIFT (Engbert et al., 2002, 2005); SEAM (Rabe et al., 2024) and the aforementioned OB1 Reader as well as the existing eye-tracking research on word transpositions.

5.1. Summary and key findings

There are four main conclusions that can be drawn from the three experiments I have presented with respect to the five computational models discussed in this thesis. Firstly, readers likely do not detect word transpositions in parafoveal vision and processing on the word(s) preceding the word transposition is not affected by their presence. Secondly, readers rapidly detect the ungrammaticality created by transposing two words upon fixating the first of the two transposed words which causes them significant and immediate disruption to processing. Third, *Transposed-Word* effects exist, however it is highly unlikely that they require the processing mechanisms proposed by the OB1 Reader model (Snell, van Leipsig, et al., 2018). Fourth, when the order of two target words changes between parafoveal and foveal vision, participants are sensitive to that change, and it disrupts their processing on the first word of the transposed pair (Target Word 1 region). Furthermore, the disruption to processing is likely driven by having an invalid preview of the Target Word 1 region alone, rather than having an invalid preview for both Target Word 1 and the second word of the transpose pair (Target Word 2 region). In the following paragraphs I will discuss each conclusion in turn and focus on the methodological and theoretical implications of these conclusions with respect to how word transpositions affect reading.

The first conclusion is supported by several findings from all three experiments and is fairly straightforward. Firstly, reading times on the Pre-Target region were not affected by the presence of a word transposition in parafoveal vision in any of the three experiments. This set of findings is also consistent with the findings by Huang and Staub (2021a) who also showed that prior to encountering an ungrammaticality, readers did not exhibit any disruption to processing. Notably, in their case, the sentences became ungrammatical at the second of the two transposed target words. Hence, the present findings on the pre-target word across all three experiments may be more comparable to the patterns of results observed by Huang and

Staub (2021a) on the first target word. Secondly, the presence of a word transposition in parafoveal vision also did not affect processing on the First Word region in Experiment 1. Finally, accuracy on the GDT in Experiment 2 was not significantly affected by the presence of a word transposition in parafoveal vision. Taken together, these findings clearly show that readers had little to no sensitivity to the ungrammaticality created by transposing the two target words in preview. Consequently, processing prior to encountering the ungrammaticality was not disrupted. This set of findings is particularly relevant and problematic for the SWIFT (Engbert et al., 2002, 2005) and SEAM (Rabe et al., 2024) models since both models predict that readers integrate (post)lexical information across words. Therefore, if processing occurs as per SWIFT and SEAM, there should have been robust parafoveal-on-foveal effects on the Pre-Target regions (and the First Word region for Experiment 1) associated with the presence of a word transposition in parafoveal vision. Instead, the lack of any such effects is consistent with the processing accounts of E-Z Reader (Reichle, et al., 1998; Reichle 2011), Über Reader (Reichle, 2021) as well as OB1 Reader (Snell, van Leipsig, et al., 2018). Consequently, these findings themselves are not sufficient to determine whether processing is serial and sequential or parallel. However, they do provide compelling evidence against the processing accounts of SWIFT and SEAM.

There are several findings across Experiment 1 and Experiment 2 in particular, which support the second conclusion. Namely, in Experiment 1 and in Experiment 2, upon fixating on the Target Word 1 region, readers experienced immediate and significant disruption to processing associated with the presence of a word transposition in foveal vision. This disruption was prolonged and still visible at the point of fixating on Target Word 2. Taken together, these effects provide compelling evidence that readers detected the ungrammaticality created by the word transposition as soon as they encountered it and that the detection resulted in substantial disruption to processing. Furthermore, these findings are particularly problematic for OB1 Reader since the model postulates that word transpositions should only have an effect on global but not local processing measures. Conversely, both the serial accounts of E-Z Reader and Über Reader as well as the parallel accounts of SWIFT and SEAM would predict the observed disruption. Note, again, that the stimuli in the present experiment were designed and pre-screened to ensure that the sentence became ungrammatical at the Target Word 1 region when a transposition was present. An ungrammaticality cue such as this is very likely to induce significant disruption as soon as it is detected. Even so, despite the clarity of the results discussed so far, it remains the case that on the basis of the evidence mentioned so far, it is not possible to discriminate unambiguously as to whether processing occurs in a serial, or a parallel manner.

Both previous research, as well as findings from Experiment 1 and Experiment 2 of my Ph.D. in particular, support the third conclusion. First, as reported in several studies, *Transposed-Word* effects do exist even when readers are presented with sentences word-by-word via a serial presentation technique (Hossain & White, 2023; Huang & Staub, 2023; Liu et al., 2022; Milledge et al., 2023; see also Mirault, Vandendaele, et al., 2022). To be clear, it is certainly the case that readers do sometimes fail to spot transposed word pairs in sentences during reading. The only exception to the aforementioned findings is the study by Snell and Melo (2024) which found no evidence of a *Transposed-Word* effect when presenting sentences word-by-word (see Section 1.7). Hence, evidence from serial presentation experiments seems, overall, to support the notion that *Transposed-Word* effects do not require that processing occurs as per OB1 Reader since they can be observed even when parallel processing is prevented, and serial processing is enforced. In other words, the fact that readers sometimes fail to detect word transpositions, is likely not due to words being processed in parallel and being assigned to positions within a sentence via a spatiotopic sentence-level mechanism. The present thesis provides further support for this conclusion in terms of two sets of key findings primarily (Experiment 1 and Experiment 2). First, as discussed above, encountering a word transposition led to significant and immediate disruption to reading visible in local eye movement measures in both Experiment 1 and Experiment 2. This is inconsistent with the prediction from the OB1 Reader model that word transpositions should only cause disruption visible in global but not local eye movement processing measures (see Snell & Grainger, 2019a).

The second set of findings pertains to the accuracy and RT results on the GDT in Experiment 1. Readers in Experiment 1 performed equally well on the GDT for sentences containing a word transposition only and sentences containing a final ungrammatical word only. This finding is inconsistent with the predictions derived from the OB1 Reader model. According to the model, sentences containing a word transposition should be more difficult to judge for grammaticality than sentences containing an ungrammatical final word. This is because, the former type of sentence can be resolved into a grammatically correct sentence by assigning words to their correct positions, while the latter type of sentence, cannot be resolved into a grammatically correct sentence regardless of what order words are assigned. This is a very strong claim of the OB1-Reader model. Importantly, this pattern of effects occurred in the context of a clear *Transposed-Word* effect such that participants were slower and made more errors on the GDT when judging sentences containing a word transposition only, versus sentences containing both a word transposition and an ungrammatical final word, consistent with previous research (e.g., Liu et al., 2020, 2021; Mirault et al., 2018, 2020). I am focusing on the presence of a *Transposed-Word* effect because one of the potential arguments as to why in Experiment 1, readers detected the ungrammaticality in sentences containing a word transposition only approximately 92% of the time, is that they were presented with 75% ungrammatical sentences, while previous studies contained 50% grammatically correct and 50% ungrammatical sentences (e.g., Hossain & White, 2023; Liu et al., 2020; Milledge et al., 2023; Mirault et al., 2018). While it is true that Experiment 1 featured more ungrammatical than grammatical sentences, there was still a robust *Transposed-Word* effect suggesting that the prevalence of ungrammatical sentences is not a strong modulator of that effect. Furthermore, one of the two published eye-tracking investigations of *Transposed-Word* effects (Huang & Staub, 2021a) also contained more ungrammatical than grammatical sentences, and in their case, there were 155 ungrammatical sentences out of 260 sentences in total, meaning that the prevalence of ungrammatical sentences was approximately 60%. Hence, the observation that participants were equally good at judging the (un)grammaticality of sentences containing a word transposition only versus sentences containing a final ungrammatical word only is also likely not due to the fact that there were 75% ungrammatical sentences in Experiment 1 (i.e., the effects are very likely not due to response bias). Instead, it seems very likely that readers simply fail to detect an ungrammaticality with increased likelihood when a “No” response is required due to a single ungrammaticality within a sentence, regardless of what that ungrammaticality is, than when they are presented with a sentence containing two ungrammaticalities. As discussed before, this is inconsistent with the OB1 Reader model account of how word transpositions are processed during reading.

On the basis of these three conclusions, it should further be evident that eye movements are an appropriate tool for investigating how word transpositions affect reading. This is an important methodological point to make, because previously Snell and Grainger (2019a) have previously argued that the effects of word transpositions on processing are not detectable via eye-tracking during sentence reading. Snell and Grainger instead suggested that word transpositions affect variables such as accuracy and RTs on the GDT which reflect global processing at the trial level. The authors suggested that local eye movement measures are not sensitive to the effects of word transpositions because readers rapidly construct a sentence-level spatiotopic representation which handles the order in which words are positioned within the sentence as per the OB1 Reader model (Snell, van Leipsig, et al., 2018). The present findings clearly show that eye movements are sensitive to the effects of word transposition on processing. This sensitivity is first observed at the point of fixating the first of the two transposed target words and is still evident at the point of fixating the second of the two transposed target words (see Section 2.4.4; Section 3.4.4). Furthermore, the interactive effects between final word grammaticality and transposition on Target Word 2 in Experiment 1 (see Section 2.4.4), show that local eye movement measures are sensitive to the comparison between sentences containing a word transposition only and sentences containing both a word transposition and an ungrammatical final word, which has been an issue that is central to multiple previous investigations of the *Transposed-Word* effect (e.g., Hossain & White, 2023; Liu et al., 2020, 2021, 2022; Milledge et al., 2023; Mirault et al., 2018, 2020). Overall, the present three experiments clearly indicate that eye movements can, and should, be utilised to investigate the effects of word transpositions on reading because they provide an insight into online processing as the word transpositions are encountered which is not possible when simply using global processing measures such as RTs and accuracy.

With respect to the fourth conclusion, the eye movement findings on Target Word 1 from both Experiment 2 and Experiment 3, suggest that processing was disrupted when the two target words were transposed in preview versus foveal vision. From Experiment 2, it is evident that early processing on Target Word 1 was disrupted by the preview change only when the two target words were in their correct order in foveal vision. Conversely, when the two target words were presented in their correct order in preview, but were transposed in foveal vision, the preview change added to the disruption associated with the presence of a foveal transposition only for later measures (see Section 3.4.4). These two findings suggest that the disruption was likely driven by the visual and/or orthographic change, rather than the syntactic mismatch between the two target words in preview versus foveal vision. Furthermore, in Experiment 3, the disruption on Target Word 1 was comparable for most measures when readers were presented with a pair of transposed target words in preview versus strings of letters that did not form valid words in place of the two targets in preview. One final piece of evidence is that disruption on Target Word 1 in Experiment 3 was comparable when readers received a masked preview of only Target Word 1 compared to when they received a masked preview of both target words. Overall, these findings provide further support for the first conclusion that readers did not detect the ungrammaticality created by transposing two target words in preview. As discussed above, the disruption to processing on Target Word 1 in both Experiment 2 and Experiment 3 associated with having a transposed target word pair in preview was primarily driven by the visual and/or orthographic mismatch between the preview at the Target Word 1 position and foveal processing of Target Word 1. These findings are clearly inconsistent with the predictions of SWIFT (Engbert et al., 2002, 2005) and SEAM (Rabe, et al., 2024) since both models predict that there should be robust parafoveal-on-foveal and parafoveal preview n+2 effects irrespective of whether readers receive a valid or invalid preview of word *n+1*.

Based on these four conclusions it should be evident that both SWIFT (Engbert et al., 2002, 2005) and SEAM (Rabe et al., 2024) provide a poor fit for the data presented across the three experiments. While OB1 Reader (Snell, van Leipsig, et al., 2018) can account for more of the present findings than SWIFT and SEAM, it still fails to account for a large proportion of findings especially with respect to how word transpositions affected local eye movement measures at the point of fixating the transposed words. The two models which provide the best fit for the present data are E-Z Reader (Reichle et al., 1998, 2011) and Über Reader (Reichle, 2021). However, it is important to be clear that it is also the case that readers sometimes *do* fail to detect word transpositions. Consequently, both serial models (E-Z

Reader and Über Reader) still need to be further expanded and modified in order to explain how this could happen under serial lexical processing constraints.

Two further pieces of evidence are also of importance to this discussion. Firstly, in Experiment 1, there was some evidence that readers were sensitive to the ungrammaticality of the final sentence word at the point of fixating on the Target Word 2 region. This should not occur if lexical processing is serial and sequential as per E-Z Reader and Über Reader. Furthermore, while both SWIFT and SEAM would predict that there should have been a (post)lexical parafoveal-on-foveal effect on Target Word 2, the presence of an ungrammatical final word should have resulted in additional disruption. Instead, the presence of an ungrammatical final word reduced reading times on Target Word 2 and this reduction seemed to be driven by the fact that readers were more likely to directly fixate the Post-Target when it was ungrammatical than grammatical. Such patterns of effects are not consistent with any of the five models and indicate that processing of line final words may not be entirely comparable to processing of line internal words. What exactly may be different, however is an empirical question that requires further research. Also, further developments in the computational models will be required to account for such processing differences (focusing on where a word within a line of text and on from which words line final fixations and return sweep launch fixations are made).

The second finding of interest is that in Experiment 3, there was a delayed parafoveal-on-foveal effect on Target Word 1 associated with having a masked preview of Target Word 2 (see Section 4.4.4). As discussed previously (see Schotter et al., 2014), it is possible for readers to obtain a parafoveal preview of word *n+2* at the point of fixating word *n*. However, this should only result in processing differences at the point of fixating word *n+2*. In Experiment 3, processing differences were observed at the point of fixating word *n+1* (Target Word 1) instead of at the point of fixating word *n+2* (Target Word 2). Notably, these processing differences were only observed when Target Word 1 was not masked in preview suggesting that readers were only able to parafoveally preview Target Word 2 when they were first able to parafoveally process (and possibly identify) Target Word 1. The fact that readers only exhibited a parafoveal sensitivity to Target Word 2 when Target Word 1 was not masked in preview is consistent with the serial assumptions of E-Z Reader and Über Reader. However, the timing of that sensitivity was earlier than expected. Consequently, future iterations of both models need to consider that effects associated with obtaining a parafoveal preview of word *n+2* may be observed earlier than might be ordinarily expected even though words are lexically processed serially and sequentially.

5.2. Future directions

As discussed above, the present thesis aimed to answer several key questions with respect to how word transpositions may impact processing during reading. That said, there are multiple outstanding questions with respect to word transpositions in particular and transpositions of linguistic units such as syllables and morphemes more generally. I will first focus on two points with respect to word transpositions and will then discuss two further points with respect to transpositions of other types of linguistic units.

The effects of word transpositions on grammaticality decisions have been fairly well documented across alphabetic spaced (e.g., English – Hossain & White, 2023; Huang & Staub, 2021a, b, 2022, 2023; Milledge et al., 2023) and logographic unspaced (e.g., Chinese, Liu et al., 2020, 2021, 2022) languages. However, the effect of word transpositions on oculomotor behaviour so far has only been examined in alphabetic spaced scripts. This is an important point to make for two reasons. First, in languages such as English and French, words are clearly demarcated by spaces and vary in length substantially. The spacing of such languages intrinsically informs readers of where one word ends and another starts, meaning that word segmentation is fairly straightforward. The spatial spread (length variability) of words, means that in a lot of scenarios, some of the letters belonging to the immediately upcoming word in the string of text may extend into the parafovea and even periphery where visual acuity is lower than in the fovea. Hence, parafoveal preprocessing of the upcoming word can be modulated by its length (e.g., see Rayner, 2009). Chinese on the other hand, is a logographic unspaced script in which a majority of words are comprised of one or two characters and consequently defining what a word is, to begin with, can be quite challenging even for experienced readers (e.g., Zang, 2019). In addition, the smaller variability in length across words in Chinese, means that in a lot of circumstances, more than one word may appear in foveal and parafoveal vision at any given fixation, meaning that processing of the upcoming word is not affected by reduced visual acuity associated with parafoveal and peripheral vision. These factors may in part explain why lexical parafoveal-on-foveal and parafoveal preview effects have been somewhat more prevalent in Chinese compared to alphabetic spaced languages. Therefore, it would be beneficial to explore the effects of word transpositions on eye movements when reading Chinese and logographic scripts in general.

In a similar manner with respect to spacing, there are multiple alphabetic languages which are either unspaced (e.g., Arabic and Hebrew e.g., Hermena et al., 2021; Velan et al., 2013) or agglutinative (e.g., Finnish). Investigating how word transpositions may affect reading as indexed by eye movements in such languages would enable researchers to better understand processing of multimorphemic words as well as words which may vary even more in length than words in alphabetic spaced languages such as English.

So far research on linguistic transpositions has primarily focused on letter and word transpositions. The importance of order for other linguistic units has received relatively little attention (see Section 1.7). For instance, no eye-tracking reading research, to my knowledge, has investigated the effects of syllable order violations on word recognition. Similarly, very few studies have investigated morpheme transpositions. Specifically, only a single eye-tracking study in English (Angele & Rayner, 2013) has explored the effects of morpheme order violations in parafoveal vision. Additionally, one study in Chinese (Gu et al., 2015) explored the effects of character transpositions in bimorphemic words in preview. Since in this study the two characters also functioned as the two morphemes of the word, it can be considered an investigation of morpheme transpositions in Chinese. Furthermore, a majority of words in Chinese are, in fact, one character long, meaning that it is not always clear how a character may be processed within text. In other words, it may be difficult to disentangle whether transposition effects in Chinese are driven by processing of characters as orthographic, morphological or lexical units. All these examples highlight the need for more research on transpositions of linguistic units other than letters and words across languages with varying properties. One final point of consideration is that studies on morpheme and character transpositions in reading have, so far, exclusively focused on parafoveal processing (e.g., Angele & Rayner, 2013; Gu et al., 2015). In contrast, the majority of published work on *Transposed-Word* effects has focused on foveally presented word transpositions. Hence it would be interesting to explore how transpositions of linguistic units beyond letters and words may impact reading, as indexed by eye movements, when they can be directly fixated, rather than appearing solely in preview.

5.3. Conclusion

The three studies presented in this thesis were formulated, designed, and conducted with the aim of providing novel and theoretically relevant insights into how reading is affected by word order violations. The current thesis provides converging evidence that words are lexically processed in a serial and sequential manner. However, the two serial computational models (E-Z Reader: Reichle et al., 1998; Reichle, 2011; and Über Reader: Reichle, 2021) are not able to explain all findings from the present three experiments and need to be further developed in order to more accurately account for how readers process text, especially with respect to how they may sometimes fail to detect the ungrammaticality created by transposing two adjacent words. Additionally, as per the example from Rayner et al. (2004; see Section 1.4.1), explaining how word transpositions affect reading via OB1 Reader is similar to *chopping carrots* with an *axe*. That is, it is highly unlikely that word transpositions (i.e., *carrots*) are processed (i.e., *chopped*) in the manner proposed by OB1 Reader (i.e., *with an axe*). Lastly, the present thesis clearly demonstrates that eye movements are an appropriate measure for examining the impact of word transpositions on reading despite the assertions by Snell and Grainger (2019a) and the OB1 Reader model (Snell, van Leipsig, et al., 2018).

Reference List

Andrews, S., & Veldre, A. (2019). What is the most plausible account of the role of parafoveal processing in reading?. *Language and Linguistics Compass, 13*(7), e12344. https://doi.org/10.1111/lnc3.12344

Angele, B., & Rayner, K. (2013). Eye movements and parafoveal preview of compound words: Does morpheme order matter? Quarterly Journal of Experimental Psychology, 66(3), 505-526. https://doi.org/10.1080/17470218.2011.644572

Angele, B., & Rayner, K. (2011). Parafoveal processing of word n + 2 during reading: Do the preceding words matter? *Journal of Experimental Psychology: Human Perception and Performance, 37*(4), 1210–1220. https://doi.org/10.1037/a0023096

Angele, B., Slattery, T. J., & Rayner, K. (2016). Two stages of parafoveal processing during reading: Evidence from A display change detection task. *Psychonomic Bulletin & Review*, *23*(4), 1241–1249. https://doi.org/10.3758/s13423-015-0995-0

Angele, B., Slattery, T. J., Yang, J., Kliegl, R., & Rayner, K. (2008). Parafoveal processing in reading: Manipulating n+1 and n+2 previews simultaneously. *Visual Cognition, 16*(6), 697–707. https://doi.org/10.1080/13506280802009704

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

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/10.1037/02787393.32.2.416

Balota, D. A., & Rayner, K. (2012). Word Recognition Processes In Foveal and Parafoveal Vision: The Range of Influence of Lexical Variables. In *Basic Processes in Reading* pp. 198–232, Routledge.

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

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*, 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

Basso, M. A., & Wurtz, R. H. (1998). Modulation of neuronal activity in superior colliculus by changes in target probability. *Journal of Neuroscience, 18*(18), 7519-7534. https://doi.org/10.1523/JNEUROSCI.18-18-07519.1998

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

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendieck, G., Green, P., & Fox, J. (2011). Package 'lme4'. *Linear mixed-effects models using S4 classes. R package version, 1*(6). https://cran.microsoft.com/snapshot/2020-04-13/web/packages/lme4/lme4.pdf

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/10.1016/00426989(79)90222-0

Bever, T. (1970). *Cognitive Basis for Linguistic Structures.* New York: Wiley Press.

Binder, K. S., Duffy, S. A., & Rayner, K. (2001). The effects of thematic fit and discourse context on syntactic ambiguity resolution. *Journal of Memory and Language, 44*(2), 297-324. https://doi.org/10.1006/jmla.2000.2754

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., 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., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2014). Reading transposed text: effects of transposed letter distance and consonant-vowel status on eye movements. *Attention Perception Psychophysics 76,* 2424–2440 https://doi.org/10.3758/s13414-014-0707-2

Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content. *Journal of psycholinguistic research*, *31*, 25-44. https://doi.org/10.1023/A:1014324220455

Bremmer, F., Kubischik, M., Hoffmann, K. P., & Krekelberg, B. (2009). Neural dynamics of saccadic suppression. *Journal of Neuroscience*, *29*(40), 12374-12383. https://doi.org/10.1523/JNEUROSCI.2908-09.2009

Brothers, T., Hoversten, L. J., & Traxler, M. J. (2017). Looking back on reading ahead: No evidence for lexical parafoveal-on-foveal effects. *Journal of Memory and Language*, *96*, 9-22. https://doi.org/10.1016/j.jml.2017.04.001

Brysbaert, M. (1996). Modifier attachment in sentence parsing: Evidence from Dutch. The *Quarterly Journal of Experimental Psychology: Section A*, *49*(3), 664-695. https://doi.org/10.1080/713755636

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In *Eye guidance in reading and scene perception* pp. 125-147. Elsevier Science Ltd. https://doi.org/10.1016/B978-008043361-5/50007-9

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Cognitive processes in eye guidance. In *Word skipping: Implications for theories of eye movement control in reading* pp. 53-78. Oxford, UK: Oxford University Press.

Büttner-Ennever, J. A., Cohen, B., Pause, M., & Fries, W. (1988). Raphe nucleus of the pons containing omnipause neurons of the oculomotor system in the monkey, and its homologue in man. *Journal of comparative neurology*, *267*(3), 307-321. https://doi.org/10.1002/cne.902670302

Carpenter, P. A., & Daneman, M. (1981). Lexical retrieval and error recovery in reading: A model based on eye fixations. *Journal of Verbal Learning and Verbal Behavior, 20*(2), 137-160. https://doi.org/10.1016/S0022-5371(81)90357-1

Chang, M., Hao, L., Zhao, S., Li, L., Paterson, K. B., & Wang, J. (2020). Flexible parafoveal encoding of character order supports word predictability effects in Chinese reading: Evidence from eye movements. Attention, Perception, & Psychophysics, 82, 2793-2801. https://doi.org/10.3758/s13414-020-02050-x

Chiu, T.-Y., & Drieghe, D. (2023). The role of visual crowding in eye movements during reading: Effects of text spacing. *Attention, Perception, & Psychophysics, 85*(8), 2834–2858. https://doi.org/10.3758/s13414-023-02787-1

Chui, T. Y., Song, H., & Burns, S. A. (2008). Adaptive-optics imaging of human cone photoreceptor distribution. *Journal of the Optical Society of America A, 25*(12), 3021. https://doi.org/10.1364/josaa.25.003021

Clifton Jr, C., Ferreira, F., Henderson, J. M., Inhoff, A. W., Liversedge, S. P., Reichle, E. D., & Schotter, E. R. (2016). Eye movements in reading and information processing: Keith Rayner's 40 year legacy. *Journal of Memory and Language, 86*, 1-19. https://doi.org/10.1016/j.jml.2015.07.004

Clifton Jr, C., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language, 49*(3), 317-334. https://doi.org/10.1016/S0749-596X(03)00070-6

Cohen, A. L., & Staub, A. (2014). Online processing of novel noun–noun compounds: Eye movement evidence. *Quarterly Journal of Experimental Psychology, 67*(1), 147-165. https://doi.org/10.1080/17470218.2013.79639

Cui, L., Drieghe, D., Yan, G., Bai, X., Chi, H., & Liversedge, S. P. (2013). Parafoveal processing across different lexical constituents in Chinese reading. *Quarterly Journal of Experimental Psychology, 66*(2), 403-416. https://doi.org/10.1080/17470218.2012.720265

Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology, 292*(4), 497-523. https://doi.org/10.1002/cne.902920402

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across fixations in reading. *The Oxford handbook of reading,* 245-260.

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2017). Reading sentences of uniform word length: Evidence for the adaptation of the preferred saccade length during reading. *Journal of Experimental Psychology: Human Perception and Performance, 43*(11), 1895–1911. https://doi.org/10.1037/xhp0000416

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2018). Reading sentences of uniform word length – II: Very rapid adaptation of the preferred saccade length. *Psychonomic Bulletin & Review, 25*(4), 1435–1440. https://doi.org/10.3758/s13423-018-1473-2

Cutter, M. G., Martin, A. E., & Sturt, P. (2020). Capitalization interacts with syntactic complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(6), 1146–1164. https://doi.org/10.1037/xlm0000780

Cutter, M. G., Martin, A. E., & Sturt, P. (2020). The activation of contextually predictable words in syntactically illegal positions. *Quarterly Journal of Experimental Psychology, 73(*9), 1423-1430. https://doi.org/10.1177/1747021820911021

Degno, F., Loberg, O., & Liversedge, S. P. (2021). Co-registration of eye movements and fixation—related potentials in natural reading: Practical issues of experimental design and data analysis. *Collabra: Psychology*, *7*(1). https://doi.org/10.1525/collabra.18032

Degno, F., Loberg, O., Zang, C., Zhang, M., Donnelly, N., & Liversedge, S. P. (2019a). A co-registration investigation of inter-word spacing and parafoveal preview: Eye movements and fixation-related potentials. *PLOS ONE, 14*(12). https://doi.org/10.1371/journal.pone.0225819

Degno, F., Loberg, O., Zang, C., Zhang, M., Donnelly, N., & Liversedge, S. P. (2019b). Parafoveal previews and lexical frequency in natural reading: Evidence from eye movements and fixation-related potentials. *Journal of Experimental Psychology: General*, *148*(3), 453–474. https://doi.org/10.1037/xge0000494

Department of Education. (2015). Reading: The next steps. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/409409/Reading_the_next_steps.pdf

Deutsch, A., & Rayner, K. (1999). Initial fixation location effects in reading Hebrew words. *Language and Cognitive Processes, 14*(4), 393-421. https://doi.org/10.1080/016909699386284

Devescovi, A., Bates, E., D'amico, S., Hernandez, A., Marangolo, P., Pizzamiglio, L., & Razzano, C. (1997). An on-line study of grammaticality judgements in normal and aphasic speakers of Italian. *Aphasiology*, *11*(6), 543-579. https://doi.org/10.1080/02687039708248490

Dienes, Z. (2008). Understanding psychology as a science: An introduction to scientific and statisticalinference. Palgrave Macmillan.

Dirix, N., & Duyck, W. (2017). An eye movement corpus study of the age-of-acquisition effect. *Psychonomic bulletin & review, 24,* 1915-1921. https://doi.org/10.3758/s13423-017-1233-8

Drieghe, D. (2008). Foveal processing and word skipping during reading. *Psychonomic Bulletin & Review, 15,* 856-860. https://doi.org/10.3758/PBR.15.4.856

Drieghe, D. (2011). Parafoveal-on-foveal effects on eye movements during reading. *Oxford Handbooks Online*. https://doi.org/10.1093/oxfordhb/9780199539789.013.0046

Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology, 16*(1-2), 79-103. https://doi.org/10.1080/09541440340000141

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

Drieghe, D., Rayner, K., & Pollatsek, A. (2008). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *Quarterly Journal of Experimental Psychology, 61*(8), 1239-1249. https://doi.org/10.1080/17470210701467953

Drieghe, D., Veldre, A., Fitzsimmons, G., Ashby, J., & Andrews, S. (2019). The influence of number of syllables on word skipping during reading revisited. *Psychonomic Bulletin & Review, 26*(2), 616–621. https://doi.org/10.3758/s13423-019-01590-0

Dufour, S., Mirault, J. and Grainger, J. (2022) Transposed-word effects in speeded grammatical decisions to sequences of spoken words. *Scientific Reports, 12*(1). https://doi.org/10.1038/s41598-022-26584-2

Dugdale, G. & Clark, C. (2008). Literacy changes lives. An advocacy resource. https://cdn.literacytrust.org.uk/media/documents/2008_09_02_free_research_-_Literacy_changes_lives_2008_exec_summary_CY4wm3v.pdf

Ehrlich, S. E, & 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., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research*, *42*(5), 621-636. https://doi.org/10.1016/S0042-6989(01)00301-7

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

Farid, M., & Grainger, J. (1996). How initial fixation position influences visual word recognition: A comparison of French and Arabic. *Brain and Language, 53*(3), 351-368. https://doi.org/10.1006/brln.1996.0053

Ferreira, F., & Qiu, Z. (2021). Predicting syntactic structure. *Brain research, 1770*, 147632. https://doi.org/10.1016/j.brainres.2021.147632

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of psycholinguistic research, 30,* 3-20. https://doi.org/10.1023/A:1005290706460

Filik, R. (2008). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition, 106*(2), 1038-1046. https://doi.org/10.1016/j.cognition.2007.04.006

Filik, R., Paterson, K. B., & Liversedge, S. P. (2005). Parsing with focus particles in context: Eye movements during the processing of relative clause ambiguities. *Journal of Memory and Language, 53*(4), 473-495. https://doi.org/10.1016/j.jml.2005.07.004

Filik, R., Paterson, K. B., & Liversedge, S. P. (2009). The influence of only and even on online semantic interpretation. *Psychonomic bulletin & review, 16,* 678-683. https://doi.org/10.3758/PBR.16.4.678

Filik, R., Paterson, K. B., & Sauermann, A. (2011). The influence of focus on eye movements during reading. In Liversedge, S., Gilchrist, I., & Everling, S. (Eds.). *The Oxford handbook of eye movements.* OUP Oxford.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences, 22*(4), 661-674. https://doi.org/10.1017/S0140525X99002150

Francis, W. N., & Kućera, H. (1982). Frequency analysis of English usage: Lexicon and grammar. Boston: Houghton Mifflin.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), Attention and performance: *Vol. 12. The psychology of reading* pp. 559–586. Hillsdale, NJ: Erlbaum.

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/10.1016/00100285(82)90008-1

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings vs. multiple senses. *Journal of Memory and Language, 29*(2), 181-200. https://doi.org/10.1016/0749-596X(90)90071-7

Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of crosslinguistic word-order variation. *Psychological science*, *24*(7), 1079-1088. https://doi.org/10.1177/0956797612463705

Goldberg, M. E., & Segraves, M. A. (1989). The visual and frontal cortices. *Reviews of oculomotor research*, *3*, 283-313. https://europepmc.org/article/med/2486326

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: a model of letter position coding. *Psychological review*, *115*(3), 577. https://doi.org/10.1037/a0012667

Grainger, J., & Van Heuven, W. J. B. (2004). Modeling Letter Position Coding in Printed Word Perception. In P. Bonin (Ed.), *Mental lexicon: "Some words to talk about words"* pp. 1–23. Nova Science Publishers.

Grainger, J., Mathôt, S., & Vitu, F. (2014). Tests of a model of multi-word reading: Effects of parafoveal flanking letters on foveal word recognition. *Acta Psychologica*, *146*, 35–40. http://dx.doi.org/10.1016/j.actpsy.2013.11.014

Gu, J., Li, X., & Liversedge, S. P. (2015). Character order processing in Chinese reading. Journal of Experimental Psychology: Human Perception and Performance, 41(1), 127–137. https://doi.org/10.1037/a0038639

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

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science*, *274*(5286), 427–430. https://doi.org/10.1126/science.274.5286.427

He, L., Ma, W., Shen, F., Wang, Y., Wu, J., Warrington, K. L., Liversedge, S. P., & Paterson, K. B. (2021). Adult age differences in parafoveal preview effects during reading: Evidence from Chinese. *Psychology and Aging, 36*(7), 822–833. https://doi.org/10.1037/pag0000639

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

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. (1993). Eye movement control during reading: Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology, 47*, 201-221. https://doi.org/10.1037/h0078814

Hermena, E. W., Juma, E. J., AlJassmi, M. (2021) Parafoveal processing of orthographic, morphological, and semantic information during reading Arabic: A boundary paradigm investigation. PLoS ONE 16(8): e0254745. https://doi.org/10.1371/journal.pone.0254745

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(1), 166–190. https://doi.org/10.1037/a0033670

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 36(5), 1150–1170. https://doi.org/10.1037/a0020233

Horn, A., Buttner-Ennever, J., Wahle, P., & Reichenberger, I. (1994). Neurotransmitter profile of saccadic omnipause neurons in nucleus raphe interpositus. *Journal of Neuroscience*, *14(*4), 2032–2046. https://doi.org/10.1523/jneurosci.14-04-02032.1994

Horwitz, G. D., & Newsome, W. T. (2001). Target selection for saccadic eye movements: direction-selective visual responses in the superior colliculus. *Journal of Neurophysiology*, *86*(5), 2527-2542. https://doi.org/10.1152/jn.2001.86.5.2527

Hossain, J., & White, A. L. (2023). The transposed word effect is consistent with serial word recognition and varies with reading speed. *Cognition*, *238*, 105512. https://doi.org/10.1016/j.cognition.2023.105512

Huang, K. J., & Staub, A. (2021a). Using eye tracking to investigate failure to notice word transpositions in reading. *Cognition*, *216*, 104846. https://doi.org/10.1016/j.cognition.2021.104846

Huang, K. J., & Staub, A. (2021b). Why do readers fail to notice word transpositions, omissions, and repetitions? A review of recent evidence and theory. *Language And Linguistics Compass*, *15*(7). https://doi.org/10.1111/lnc3.12434

Huang, K. J., & Staub, A. (2022). Readers do not strongly rely on full-context information, but might utilize local word statistics, when 'correcting' word transposition errors in text. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 44. https://escholarship.org/uc/item/3xs699w8

Huang, K. J., & Staub, A. (2023). The transposed-word effect does not require parallel word processing: Failure to notice transpositions with serial presentation of words.

Psychonomic Bulletin & Review, *30*, 393-400. https://doi.org/10.3758/s13423-022-02150-9

Huey, E. B. (1908). *The Psychology and Pedagogy of Reading*. New York: The Macmillan Company.

Hyönä, J. (1993). Effects of thematic and lexical priming on readers' eye movements. *Scandinavian Journal of Psychology, 34*(4), 293–304. https://doi.org/10.1111/j.1467-9450.1993.tb01126.x

Hyönä, J. (1995). Do irregular letter combinations attract readers' attention? Evidence from fixation locations in words. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 68- 81. https://doi.org/10.1037/0096-1523.21.1.68

Hyönä, J., & Olson, R. K. (1995). Eye movement 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

Hyönä, J., Niemi, P., & Underwood, G. (1989). Reading long words embedded in sentences: Informativeness of word halves affects eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 142–152. https://doi.org/10.1037/0096-1523.15.1.142

Hyönä, J., Yan, M., & Vainio, S. (2018). Morphological structure influences the initial landing position in words during reading Finnish. *Quarterly Journal of Experimental Psychology*, *71*(1), 122-130. https://doi.org/10.1080/17470218.2016.1267233

Hyvärinen, J. (1982). Posterior parietal lobe of the primate brain. *Physiological Reviews, 62(*3), 1060–1129. https://doi.org/10.1152/physrev.1982.62.3.1060

Inhoff, A. W. (1984). Two stages of word processing during eye fixations in the reading of prose. *Journal of Verbal Learning and Verbal Behavior, 23*, 612-624. https://doi.org/10.1016/S0022-5371(84)90382-7

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*(6), 431-439. https://doi.org/10.3758/BF03208203

Inhoff, A. W., Kim, A., & Radach, R. (2019). Regressions during reading. *Vision, 3(*3), 35-49. https://doi.org/10.3390/vision3030035

Inhoff, A. W., Pollatsek, A., Posner, M. I., & Rayner, K. (1989). Covert attention and eye movements during reading. *Quarterly Journal of Experimental Psychology, 41A*, 63–89. https://doi.org/10.1080/14640748908402353

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

Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In *Reading as a Perceptual Process*, pp 221–246. https://doi.org/10.1016/b978-008043642-5/50012-7

Irwin, D. E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology*, *36*(1), 1–27. https://doi.org/10.1006/cogp.1998.0682

Irwin, D. E., Carlson-Radvansky, L. A., & Andrews, R. V. (1995). Information processing during saccadic eye movements. *Acta Psychologica*, *90*(1–3), 261–273. https://doi.org/10.1016/0001-6918(95)00024-o

Jackendoff, R. S. (1972). *Semantic interpretation in generative grammar.* Cambridge, MA: MIT Press

Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, *94*, 316-339. https://doi.org/10.1016/j.jml.2017.01.004

Jäger, L. A., Mertzen, D., Van Dyke, J. A., & Vasishth, S. (2020). Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language*, *111*, 104063. https://doi.org/10.1016/j.jml.2019.104063

Johnson, R. L. (2007). The flexibility of letter coding: Nonadjacent letter transposition effects in the parafovea. In Eye Movements (pp. 425-440). Elsevier. https://doi.org/10.1016/B978-008044980-7/50021-5

Johnson, R. L. (2009). The quiet clam is quite calm: transposed-letter neighborhood effects on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 943-969. https://doi.org/10.1037/a0015572

Johnson, R. L., & Starr, E. L. (2018). The preferred viewing location in top-to-bottom sentence reading. *Quarterly Journal of Experimental Psychology, 71*(1), 220-228. https://doi.org/10.1080/17470218.2017.1307860

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

Joseph, H. S., Wonnacott, E., Forbes, P., & Nation, K. (2014). Becoming a written word: Eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading. *Cognition, 133*(1), 238-248. https://doi.org/10.1016/j.cognition.2014.06.015

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*(7–8), 846–863. https://doi.org/10.1080/13506280544000075

Juhasz, B. J., Inhoff, A. W., & Rayner, K. (2005). The role of interword spaces in the processing of English compound words. *Language and Cognitive Processes*, *20*(1–2), 291–316. https://doi.org/10.1080/01690960444000133

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

Kapoula, Z. (1985). Evidence for a range effect in the saccadic system. *Vision research, 25*(8), 1155-1157. https://doi.org/10.1016/0042-6989(85)90105-1

Kapoula, Z., & Robinson, D. A. (1986). Saccadic undershoot is not inevitable: Saccades can be accurate. *Vision research, 26*(5), 735-743. https://doi.org/10.1016/0042-6989(86)90087-8

Kemper, S., & Liu, C.-J. (2007). Eye movements of young and older adults during reading. *Psychology and Aging*, *22*(1), 84–93. https://doi.org/10.1037/0882-7974.22.1.84

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, *45*(2), 153–168. https://doi.org/10.1016/j.visres.2004.07.037

Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 68–81. https://doi.org/10.1037/0278-7393.21.1.68

Kerr, P. W. (1992). *Eye movement control during reading: The selection of where to send the eyes*. University of Illinois at Urbana-Champaign.

Kirsch, I., De Jong, J., LaFontaine, D., McQueen, J., Mendelovits, J. & Monseur, C. (2003). Reading for change: Performance and engagement across countries: Results of PISA 2000. https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33690904.pdf

Kliegl, R. (2007). Toward a perceptual-span theory of distributed processing in reading: A reply to Rayner, Pollatsek, Drieghe, Slattery, and Reichle (2007). *Journal of Experimental Psychology: General*, *136*(3), 530–537. https://doi.org/10.1037/0096-3445.136.3.530

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European journal of cognitive psychology*, *16*(1-2), 262-284. https://doi.org/10.1080/09541440340000213

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, *135*(1), 12–35. https://doi.org/10.1037/0096-3445.135.1.12

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word n+ 2. *Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1250-1255. https://doi.org/10.1037/0096-1523.33.5.1250

Leigh, R. J., & Zee, D. S. (2006). *The neurology of Eye Movements*. Oxford University Press.

Lettvin, J. Y. (1976). On seeing sidelong. *The Sciences*, *16*(4), 10-20. http://jerome.lettvin.com/jerome/OnSeeingSidelong.pdf

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. Cognitive Science, 29(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Li, N., Sun, D., & Wang, S. (2023). Semantic preview effect of relatedness and plausibility in reading Chinese: evidence from high constraint sentences. *Reading and Writing, 36*(5), 1319-1338. https://doi.org/10.1007/s11145-022-10350-x

Li, X., Bicknell, K., Liu, P., Wei, W., & Rayner, K. (2014). Reading is fundamentally similar across disparate writing systems: A systematic characterization of how words and characters influence eye movements in Chinese reading. *Journal of Experimental Psychology: General*, *143*(2), 895–913. https://doi.org/10.1037/a0033580

Li, X., Huang, L., Yao, P., & Hyönä, J. (2022). Universal and specific reading mechanisms across different writing systems. *Nature Reviews Psychology, 1*(3), 133-144. https://doi.org/10.1038/s44159-022-00022-6

Liu, Y., Reichle, E. D., & Li, X. (2016). The effect of word frequency and parafoveal preview on saccade length during the reading of Chinese. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 1008–1025. https://doi.org/10.1037/xhp0000190

Liu, Z., Li, Y., Cutter, M. G., Paterson, K. B., & Wang, J. (2022). A transposed-word effect across space and time: Evidence from Chinese. *Cognition*, *218*, 104922. https://doi.org/10.1016/j.cognition.2021.104922

Liu, Z., Li, Y., Paterson, K. B., & Wang, J. (2020). A transposed-word effect in Chinese reading. *Attention, Perception, & Psychophysics*, *82*(8), 3788-3794. https://doi.org/10.3758/s13414-020-02114-y

Liu, Z., Li, Y., & Wang, J. (2021) Context but not reading speed modulates transposed-word effects in Chinese reading. *Acta Psychologica, 215*, 103272. https://doi.org/10.1016/j.actpsy.2021.103272

Liu, Z., Li, Y., & Wang, J. (2024). Exploring the flexibility of word position encoding in Chinese reading: the role of transposition effects. *Language, Cognition and Neuroscience*, 1-7. https://doi.org/10.1080/23273798.2024.2417430

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, *4*(1), 6–14. https://doi.org/10.1016/s1364-6613(99)01418-7

Liversedge, S. P., Paterson, K. B., & Clayes, E. L. (2002). The influence of only on syntactic processing of "long" relative clause sentences. *The Quarterly Journal of Experimental Psychology Section A*, *55*(1), 225-240. https://doi.org/10.1080/02724980143000253

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

Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during Chinese reading. *Visual Cognition*, *22(*3-4), 441-457. https://doi.org/10.1080/13506285.2014.889260

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in psychology*, *6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171

Lu, C. C., Bates, E., Li, P., Tzeng, O., Hung, D., Tsai, C. H., Lee, S., & Chung, Y. M. (2000). Judgements of grammaticality in aphasia: The special case of Chinese. *Aphasiology*, *14*(10), 1021-1054. https://doi.org/10.1080/02687030050156593

Martinez, R., & Fernandez, A. (2010). The Social and economic impact of illiteracy: analytical model and pilot study. Unesdoc.unesco.org. https://unesdoc.unesco.org/ark:/48223/pf0000190571

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, *5*(3), 229–240. https://doi.org/10.1038/nrn1348

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

Maurice, D. M. (1984). The cornea and Sclera. *The Eye*, 1–158. https://doi.org/10.1016/b978-0-12-206921-5.50006-4

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., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the psychonomic society, 8*(5), 365-368. https://doi.org/10.3758/BF03335168

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision research, 28*(10), 1107-1118. https://doi.org/10.1016/0042-6989(88)90137-X

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception & Psychophysics, 46*(3), 245-253. https://doi.org/10.3758/BF03208086

Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2009). Parafoveal magnification: Visual acuity does not modulate the perceptual span in reading. *Psychological Science, 20*(6), 721-728. https://doi.org/10.1111/j.1467-9280.2009.02364.x

Milledge, S. V., Blythe, H. I., & Liversedge, S. P. (2021). Parafoveal pre-processing in children reading English: The importance of external letters. *Psychonomic Bulletin & Review, 28*(1), 197-208. https://doi.org/10.3758/s13423-020-01806-8

Milledge, S. V., Liversedge, S. P., & Blythe, H. I. (2022). The importance of the first letter in children’s parafoveal preprocessing in English: Is it phonologically or orthographically driven? *Journal of Experimental Psychology: Human Perception and Performance, 48*(5), 427–442. https://doi.org/10.1037/xhp0000998

Milledge, S. V., Zang, C., Liversedge, S. P., & Blythe, H. I. (2022). Phonological parafoveal pre-processing in children reading English sentences. *Cognition, 225,* 105141. https://doi.org/10.1016/j.cognition.2022.105141

Milledge, S.V., Bhatia, N., Mensah-Mcleod, L., Raghvani, P., McGowan, V. A., Elsherif, M. M., Cutter, M. G., Wang, J., Liu, Z., & Patterson, K. B., (2023). The transposed-word effect provides no unequivocal evidence for parallel processing. *Attention Perception Psychophysics*. https://doi.org/10.3758/s13414-023-02721-5

Mirault, J., Declerck, M., & Grainger, J. (2022). Fast priming of grammatical decisions: Repetition and transposed-word priming effects. *The Royal Society Open Science 9*, 211082 https://doi.org/10.1098/rsos.211082

Mirault, J., Guerre-Genton, A., Dufau, S., & Grainger, J. (2020). Using virtual reality to study reading: An eye-tracking investigation of transposed-word effects. *Methods In Psychology*, *3*, 100029. https://doi.org/10.1016/j.metip.2020.100029

Mirault, J., Leflaëc, C. & Grainger, J. (2022) Orthographic relatedness and transposed-word effects in the grammatical decision task. *Attention, Perception, &; Psychophysics*, *84*(3), 1043–1051. https://doi.org/10.3758/s13414-021-02421-y

Mirault, J., Snell, J., & Grainger, J. (2018). You That Read Wrong Again! A Transposed-Word Effect in Grammaticality Judgments. *Psychological Science*, *29*(12), 1922-1929. https://doi.org/10.1177/0956797618806296

Mirault, J., Vandendaele, A., Pegado, F., & Grainger, J. (2022) Transposed-word effects when reading serially. *PLOS ONE*, *17*(11). https://doi.org/10.1371/journal.pone.0277116

Mirault, J., Vandendaele, A., Pegado, F., & Grainger, J. (2023). The impact of atypical text presentation on transposed-word effects. *Attention, Perception, & Psychophysics*, 1-10. https://doi.org/10.3758/s13414-023-02760-y

Molday, R. S., & Moritz, O. L. (2015). Photoreceptors at a glance. *Journal of Cell Science*, *128*(22), 4039–4045. https://doi.org/10.1242/jcs.175687

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. https://doi.org/10.1037/00961523.16.2.268

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

Müller, L. J., Marfurt, C. F., Kruse, F., & Tervo, T. M. T. (2003). Corneal nerves: Structure, contents and function. *Experimental Eye Research, 76*(5), 521–542. https://doi.org/10.1016/s0014-4835(03)00050-2

Munoz, D. P. (2002). Commentary: Saccadic Eye Movements: Overview of Neural Circuitry. *The Brain's Eye: Neurobiological and Clinical Aspects of Oculomotor Research*, 89–96. https://doi.org/10.1016/s0079-6123(02)40044-1

Munoz, D. P. and Wurtz, R.H. (1993a) Fixation cells in monkey superior colliculus: I. characteristics of cell discharge. *Journal of neurophysiology. 70*, 559–575 https://doi.org/10.1152/jn.1993.70.2.559

Munoz, D. P. and Wurtz, R.H. (1993b) Fixation cells in monkey superior colliculus: II. reversible activation and deactivation. *Journal of neurophysiology. 70*, 576–589 https://doi.org/10.1152/jn.1993.70.2.576

Munoz, D. P., & Wurtz, R. H. (1992). Role of the rostral superior colliculus in active visual fixation and execution of express saccades. *Journal of Neurophysiology*, *67*(4), 1000–1002. https://doi.org/10.1152/jn.1992.67.4.1000

Munoz, D. P., & Wurtz, R. H. (1995). Saccade-related activity in monkey superior colliculus. II. Spread of activity during saccades. *Journal of neurophysiology*, *73*(6), 2334-2348. https://doi.org/10.1152/jn.1995.73.6.2334

Nazir, T. A. (1991). On the role of refixations in letter strings: The influence of oculomotor factors. *Perception & Psychophysics, 49*(4), 373–389. https://doi.org/10.3758/bf03205995

Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and Peripheral Vision. *Vision Research*, *32*(4), 771–777. https://doi.org/10.1016/0042-6989(92)90192-l

Ni, W. (1996). Sidestepping garden paths: Assessing the contributions of syntax, semantics and plausibility in resolving ambiguities. *Language and Cognitive Processes, 11*(3), 283–334. https://doi.org/10.1080/016909696387196

Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. (1998). Anomaly detection: Eye movement patterns. *Journal of Psycholinguistic Research, 27,* 515-539. https://doi.org/10.1023/A:1024996828734

Nuthmann, A., & Kliegl, R. (2009). Preferred viewing locations: A validation and an extension. *Perception, 38*(6), 901-902. https://pubmed.ncbi.nlm.nih.gov/19806983/

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision research, 45*(17), 2201-2217. https://doi.org/10.1016/j.visres.2005.02.014

O'Regan J. K. (1981). The convenient viewing position hypothesis. In Fisher D. F., Monty R. A., Senders J. W. (Eds.), *Eye movements: Cognition and visual perception* pp. 289–298. Hillsdale, NJ: Lawrence Erlbaum.

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

O’Regan, 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

Osaka, N., & Oda, K. (1991). Effective visual field size necessary for vertical reading during Japanese text processing. *Bulletin of the Psychonomic Society*, *29*(4), 345-347. https://doi.org/10.3758/BF03333939

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, 47*(7), 1186–1203. https://doi.org/10.1037/xlm0000989

Pan, J., & Yan, M. (2024). The perceptual span in traditional Chinese. Language and *Cognition, 16*(1), 134-147. https://doi.org/10.1017/langcog.2023.24

Parsons, S., & Bynner, J. (2008). New light on literacy and numeracy in Scotland. Edinburgh: Scottish Government. https://www.gov.scot/binaries/content/documents/govscot/publications/research-and-analysis/2008/01/new-light-adult-literacy-numeracy-scotland-evidence-2004-survey-1970-british-cohort-study-bcs70/documents/0055503-pdf/0055503-pdf/govscot%3Adocument/0055503.pdf

Paterson, K. B., & Jordan, T. R. (2010). Effects of increased letter spacing on word identification and eye guidance during reading. *Memory & Cognition, 38*(4), 502-512. https://doi.org/10.3758/MC.38.4.502

Paterson, K. B., Almabruk, A. A., McGowan, V. A., White, S. J., & Jordan, T. R. (2015). Effects of word length on eye movement control: The evidence from Arabic.

Psychonomic bulletin & review, 22, 1443-1450. https://doi.org/10.3758/s13423-015-0809-4

Paterson, K. B., Liversedge, S. P., Filik, R., Juhasz, B. J., White, S. J., & Rayner, K. (2007). Focus identification during sentence comprehension: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, *60*(10), 1423-1445. https://doi.org/10.1080/17470210601100563

Paterson, K.B., McGowan, V.A., Warrington, K.L., Li, L., Li S., Xie, F., Chang, M., Zhao, S., Pagán, A., White, S.J., & Wang, J. (2020). Effects of normative aging on eye movements during reading. *Vision, 4*(1), 7; https://doi.org/10.3390/vision4010007.

Patson, N. D., & Warren, T. (2010). Eye movements when reading implausible sentences: Investigating potential structural influences on semantic integration. *Quarterly Journal of Experimental Psychology, 63*(8), 1516-1532. https://doi.org/10.1080/17470210903380822

Pegado, F., & Grainger, J. (2019) Dissociating lexical and sublexical contributions to transposed-word effects, *Acta Psychologica*, *201*, 102943. https://doi.org/10.1016/j.actpsy.2019.102943

Pegado, F., & Grainger, J. (2020). A transposed-word effect in same-different judgments to sequences of words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(7), 1364–1371. https://doi.org/10.1037/xlm0000776

Pegado, F., & Grainger, J. (2020). On the noisy spatiotopic encoding of word positions during reading: Evidence from the change-detection task. *Psychonomic Bulletin & Review*, *28*(1), 189–196. https://doi.org/10.3758/s13423-020-01819-3

Pickering, M. J., & Traxler, M. J. (2003). Evidence against the use of subcategorisation frequency in the processing of unbounded dependencies. *Language and Cognitive Processes, 18*(4), 469-503. https://doi.org/10.1080/01690960344000017

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., Juhasz, B. J., Reichle, E. D., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: A reversed delayed effect of word length. *Journal of Experimental Psychology: Human Perception and Performance, 34*(3), 726–750. https://doi.org/10.1037/0096-1523.34.3.726

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

Pynte, J., & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. *Vision Research*, *46*(22), 3786-3801. https://doi.org/10.1016/j.visres.2006.07.004

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rabe, M. M., Paape, D., Mertzen, D., Vasishth, S., & Engbert, R. (2024). SEAM: An integrated activation-coupled model of sentence processing and eye movements in reading. *Journal of Memory and Language, 135*, 104496. https://doi.org/10.1016/j.jml.2023.104496

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *Quarterly Journal of Experimental Psychology, 66*(3), 429-452. https://doi.org/10.1080/17470218.2012.750676

Radach, R., & McConkie, G. W. (1998). Determinants of fixation positions in words during reading. In *Eye guidance in reading and scene perception* pp. 77-100. Elsevier Science Ltd. https://doi.org/10.1016/S0042-6989(01)00166-3

Radach, R., Inhoff, A. W., Glover, L., & Vorstius, C. (2013). Contextual constraint and N+ 2 preview effects in reading. *Quarterly Journal of Experimental Psychology, 66*(3), 619-633. https://doi.org/10.1080/17470218.2012.761256

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology, 49(*2), 151–173. https://doi.org/10.1037/1196- 1961.49.2.151

Rayner, K, Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics, 31*(6), 537–550. https://doi.org/10.3758/bf03204186

Rayner, K, Yang, J., Schuett, S., & Slattery, T. J. (2013). Eye movements of older and younger readers when reading unspaced text. *Experimental Psychology*, *60*(5), 354–361. https://doi.org/10.1027/1618-3169/a000207

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

Rayner, K. (1977). Visual attention in reading: Eye Movements Reflect Cognitive Processes. *Memory & Cognition*, *5*(4), 443–448. https://doi.org/10.3758/bf03197383

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*, 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), 1457–506. https://doi.org/10.1080/17470210902816461

Rayner, K. (2014). The gaze-contingent moving window in reading: Development and review. *Visual Cognition, 22*(3-4), 242-258. https://doi.org/10.1080/13506285.2013.879084

Rayner, K., & Duffy, S.A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity and lexical ambiguity. *Memory & Cognition, 14*(3), 191-201. https://doi.org/10.3758/BF03197692

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58*(5), 734-747. https://doi.org/10.3758/BF03213106

Rayner, K., & Frazier, L. (1987). Parsing temporarily ambiguous complements. The *Quarterly Journal of Experimental Psychology, 39*(4), 657-673. https://doi.org/10.1080/14640748708401808

Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements?. *Vision research, 16*(8), 829-837. https://doi.org/10.1016/0042-6989(76)90143-7

Rayner, K., & Pollatsek, A. (1996). Reading unspaced text is not easy: Comments on the implications of Epelboim et al.'s (1994) study for models of eye movement control in reading. *Vision research, 36*(3), 461-465. https://doi.org/10.1016/0042-6989(95)00132-8

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review, 3*(2), 245-248. https://doi.org/10.3758/BF03212426

Rayner, K., & Reichle, E. D. (2010). Models of the reading process. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 787-799. https://doi.org/10.1002/wcs.68

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., Angele, B., Schotter, E. R., & Bicknell, K. (2013). On the processing of canonical word order during eye fixations in reading: Do readers process transposed word previews?. *Visual Cognition*, *21*(3), 353-381. https://doi.org/10.1080/13506285.2013.791739

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of verbal learning and verbal behavior*, *22*(3), 358-374. https://doi.org/10.1016/S0022-5371(83)90236-0

Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements and the perceptual span in older and younger readers. *Psychology and Aging*, *24*(3), 755–760. https://doi.org/10.1037/a0014300

Rayner, K., Castelhano, M. S., & Yang, J. (2010). Preview benefit during eye fixations in reading for older and younger readers. *Psychology and Aging*, *25*(3), 714–718. https://doi.org/10.1037/a0019199

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific studies of reading*, *10*(3), 241-255. https://doi.org/10.1207/s1532799xssr1003_3

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., Garrod, S., & Perfetti, C. A. (1992). Discourse influences during parsing are delayed. *Cognition*, *45*(2), 109-139. https://doi.org/10.1016/0010-0277(92)90026-E

Rayner, K., Juhasz, B. J., & Brown, S. J. (2007). Do readers obtain preview benefit from word n + 2? A test of serial attention shift versus distributed lexical processing models of eye movement control in reading. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 230–245. https://doi.org/10.1037/0096-1523.33.1.230

Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review, 12,* 1089-1093. https://doi.org/10.3758/BF03206448

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., & Ehrlich, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 529–544. https://doi.org/10.1037/0096-1523.4.4.529

Rayner, K., Pollatsek, A., Ashby, J., & Clifton Jr, C. (2012). *Psychology of reading.* Psychology Press.

Rayner, K., Raney, G. E., & Pollatsek, A. (1995). Eye movements and discourse processing. In R. F. Lorch, Jr. & E. J. O'Brien (Eds.), *Sources of coherence in reading* pp. 9–35. Lawrence Erlbaum Associates, Inc.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging, 21*(3), 448–465. https://doi.org/10.1037/0882-7974.21.3.448

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22*(5), 1188–1200. https://doi.org/10.1037/0096-1523.22.5.1188

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.

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., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics, 27*(6), 537-544. https://doi.org/10.3758/BF03198682

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

Rayner, K., Yang, J., Castelhano, M. S., & Liversedge, S. P. (2011). Eye movements of older and younger readers when reading disappearing text. *Psychology and aging, 26*(1), 214. https://doi.org/10.1037/a0021279

Reichle, E. D. (2011). Serial-attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 767–786). Oxford University Press.

Reichle, E. D. (2021). Computational models of reading: A handbook. Oxford University Press. https://doi.org/10.1093/oso/9780195370669.001.0001

Reichle, E. D., & Sheridan, H. (2015). E-Z Reader: An overview of the model and two recent applications. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 277–290). Oxford University Press.

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in cognitive sciences, 13*(3), 115-119. https://doi.org/10.1016/j.tics.2008.12.002

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

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16*(1), 1–21. https://doi.org/10.3758/pbr.16.1.1

Risse, S., & Kliegl, R. (2012). Evidence for delayed parafoveal-on-foveal effects from word n+2 in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 1026–1042. https://doi.org/10.1037/a0027735

Roman, G., & Pavard, B. (1987). A comparative study: How we read in Arabic and French. In *Eye movements from physiology to cognition* pp. 431-440. Elsevier.

Rosenholtz, R. (2016). Capabilities and limitations of Peripheral Vision. *Annual Review of Vision Science*, *2*(1), 437–457. https://doi.org/10.1146/annurev-vision-082114-035733

Saenger, P. (1997). *Space between words: The origins of Silent reading.* Stanford University Press.

Sakata, H., Shibutani, H., & Kawano, K. (1980). Spatial properties of visual fixation neurons in posterior parietal association cortex of the monkey. *Journal of Neurophysiology, 43*(6), 1654–1672. https://doi.org/10.1152/jn.1980.43.6.1654

Schad, D. J., & Engbert, R. (2012). The zoom lens of attention: Simulating shuffled versus normal text reading using the SWIFT model. *Visual Cognition, 20*(4-5), 391-421. https://doi.org/10.1080/13506285.2012.670143

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & cognition, 26*(6), 1270-1281. https://doi.org/10.3758/BF03201199

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language*, *69*(4), 619-633. https://doi.org/10.1016/j.jml.2013.09.002

Schotter, E. R., & Jia, A. (2016). Semantic and plausibility preview benefit effects in English: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(*12), 1839–1866. https://doi.org/10.1037/xlm0000281

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

Schotter, E. R., Lee, M., Reiderman, M., & Rayner, K. (2015). The effect of contextual constraint on parafoveal processing in reading. *Journal of Memory and Language*, *83,* 118-139. https://doi.org/10.1016/j.jml.2015.04.005

Schotter, E. R., Reichle, E. D., & Rayner, K. (2014). Rethinking parafoveal processing in reading: Serial-attention models can explain semantic preview benefit and N+ 2 preview effects. Visual *Cognition*, *22*(3-4), 309-333. https://doi.org/10.1080/13506285.2013.873508

Schustack, M. W., Ehrlich, S. F., & Rayner, K. (1987). Local and global sources of contextual facilitation in reading. *Journal of Memory and language*, *26(*3), 322-340. https://doi.org/10.1016/0749-596X(87)90117-3

Scudder, C. A., Fuchs, A. F., & Langer, T. P. (1988). Characteristics and functional identification of saccadic inhibitory burst neurons in the alert monkey. *Journal of neurophysiology*, *59*(5), 1430-1454. https://doi.org/10.1152/jn.1988.59.5.1430

Sears, C. R., Campbell, C. R., & Lupker, S. J. (2006). Is there a neighborhood frequency effect in English? Evidence from reading and lexical decision. *Journal of*

Experimental Psychology: Human Perception and Performance, 32(4), 1040-1062. https://doi.org/10.1037/0096-1523.32.4.1040

Segraves, M. A. (1992). Activity of monkey frontal eye field neurons projecting to oculomotor regions of the Pons. *Journal of Neurophysiology, 68*(6), 1967–1985. https://doi.org/10.1152/jn.1992.68.6.1967

Shimron, J. (2006). Reading Hebrew: The language and the psychology of reading it. Routledge. https://doi.org/10.4324/9781410617187

Silvey, C., Dienes, Z., & Wonnacott, E. (2024) Bayes factors for logistic (mixed-effect) models. *Psychological Methods.* Advance online publication. https://doi.org/10.1037/met0000714

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

Smith, N. J., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition, 128*(3), 302-319. https://doi.org/10.1016/j.cognition.2013.02.013

Snell, J., & Grainger, J. (2019a). Readers are parallel processors. *Trends in Cognitive Sciences, 23*(7), 537-546. https://doi.org/10.1016/j.tics.2019.04.006

Snell, J., & Grainger, J. (2019b). Word position coding in reading is noisy. *Psychonomic Bulletin & Review, 26*(2), 609-615. https://doi.org/10.3758/s13423-019-01574-0

Snell, J., & Melo, A. N. (2024). Do Love You Me? Failure to Notice Word Transpositions is Induced by Parallel Word Processing. *Journal of Cognition, 7*(1). https://doi.org/10.5334/joc.335

Snell, J., Meeter, M., & Grainger, J. (2017). Evidence for simultaneous syntactic processing of multiple words during reading. *PLOS ONE*, *12*(3). https://doi.org/10.1371/journal.pone.0173720

Snell, J., Mirault, J., Theeuwes, J., & Grainger, J. (2023). Readers use word length information to determine word order. *Journal of Experimental Psychology: Human Perception and Performance, 49(6)*, 753–758. https://doi.org/10.1037/xhp0001107

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

Snell, J., Yeaton, J., Mirault, J., & Grainger, J. (2023). Parallel word reading revealed by fixation-related brain potentials. *Cortex*, *162*, 1–11. https://doi.org/10.1016/j.cortex.2023.02.004

Song, H., Chui, T. Y., Zhong, Z., Elsner, A. E., & Burns, S. A. (2011). Variation of cone photoreceptor packing density with retinal eccentricity and age. *Investigative Ophthalmology & Visual Science*, *52*(10), 7376. https://doi.org/10.1167/iovs.11-7199

Sperlich, A., Meixner, J., & Laubrock, J. (2016). Development of the perceptual span in reading: A longitudinal study. *Journal of experimental child psychology*, *146*, 181-201. https://doi.org/10.1016/j.jecp.2016.02.007

Spinelli, G., Yang, H., & Lupker, S. J. (2024). Rapid Serial Visual Presentation of transposed-word sequences in the grammatical decision task: an examination of the roles of temporal and spatial cues to word order. *Language, Cognition and Neuroscience*, *39*(6), 773-792. https://doi.org/10.1080/23273798.2024.2360168

Starr, M., & Inhoff, A. (2004). Attention allocation to the right and left of a fixated word: Use of orthographic information from multiple words during reading. *European Journal of Cognitive Psychology, 16*(1-2), 203-225. https://doi.org/10.1080/09541440340000150

Staub, A. (2015). The effect of lexical predictability on eye movements in reading: Critical review and theoretical interpretation. *Language and linguistics compass, 9(*8), 311-327. https://doi.org/10.1111/lnc3.12151

Staub, A., & Clifton Jr, C. (2006). Syntactic prediction in language comprehension: evidence from either... or. Journal of experimental psychology: *Learning, memory, and cognition, 32*(2), 425. https://doi.org/10.1037/0278-7393.32.2.425

Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., & Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: evidence from noun-noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 1162. https://doi.org/10.1037/0278-7393.33.6.1162

Strassman, A., Highstein, S. M., & McCrea, R. A. (1986). Anatomy and physiology of saccadic burst neurons in the alert squirrel monkey. I. Excitatory burst neurons. *Journal of Comparative Neurology, 249*(3), 337-357. https://doi.org/10.1002/cne.902490303

Strassman, A., Highstein, S. M., & McCrea, R. A. (1986). Anatomy and physiology of saccadic burst neurons in the alert squirrel monkey. II. Inhibitory burst neurons. *Journal of Comparative Neurology, 249*(3), 337-357. https://doi.org/10.1002/cne.902490304

Tiffin-Richards, S. P. (2024). The effect of word transpositions on grammaticality judgements in first and second language sentence reading. *Quarterly Journal of*

Experimental Psychology, *77(1)*. 204-216 https://doi.org/10.1177/17470218231161433

Tiffin-Richards, S. P., & Schroeder, S. (2015). 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

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of memory and language*, *47*(1), 69-90. https://doi.org/10.1006/jmla.2001.2836

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*(3), 285-318. https://doi.org/10.1006/jmla.1994.1014

Tsai, J. L., Kliegl, R., & Yan, M. (2012). Parafoveal semantic information extraction in traditional Chinese reading. *Acta Psychologica*, *141*(1), 17-23. https://doi.org/10.1016/j.actpsy.2012.06.004

van Der Linden, L., & Vitu, F. (2016). On the optimal viewing position for object processing. *Attention, Perception, & Psychophysics*, *78*, 602-617. https://doi.org/10.3758/s13414-015-1025-z

Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, *49*(3), 285-316. https://doi.org/10.1016/S0749-596X(03)00081-0

Van Gisbergen, J. A., Robinson, D. A., & Gielen, S. (1981). A quantitative analysis of generation of saccadic eye movements by burst neurons. *Journal of Neurophysiology*, *45*(3), 417–442. https://doi.org/10.1152/jn.1981.45.3.417

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly journal of experimental psychology*, *67*(6), 1176-1190. https://doi.org/10.1080/17470218.2013.850521

van Zanden, J. L., Baten, J., d'Ercole, M. M., Rijpma, A., & Timmer, M. P. (2014). How was life? OECD. https://doi.org/10.1787/9789264214262-en

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*(3), 666-689. https://doi.org/10.3758/s13423-016-1147-x

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/s13423-016-1147-x

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

Vasishth, S., Brüssow, S., Lewis, R. L., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science*, *32*, 685–712. https://doi.org/10.1080/03640210802066865

Velan, H., Deutsch, A., & Frost, R. (2013). The Flexibility of Letter-Position Flexibility: Evidence From Eye Movements in Reading Hebrew. *Journal of experimental psychology. Human perception and performance*, *39*, 1143-1152. https://doi.org/10.1037/a0031075

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. (2016a). Is semantic preview benefit due to relatedness or plausibility? *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 939–952. https://doi.org/10.1037/xhp0000200

Veldre, A., & Andrews, S. (2016b). Semantic preview benefit in English: Individual differences in the extraction and use of parafoveal semantic information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(6), 837–854. https://doi.org/10.1037/xlm0000212

Veldre, A., & Andrews, S. (2017). Parafoveal preview benefit in sentence reading: Independent effects of plausibility and orthographic relatedness. *Psychonomic Bulletin & Review*, *24*, 519-528. https://doi.org/10.3758/s13423-016-1120-8

Veldre, A., & Andrews, S. (2018a). Beyond cloze probability: Parafoveal processing of semantic and syntactic information during reading. *Journal of Memory and Language*, *100*, 1–17. https://doi.org/10.1016/j.jml.2017.12.002

Veldre, A., & Andrews, S. (2018b). 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

Veldre, A., & Andrews, S. (2018c). Parafoveal preview effects depend on both preview plausibility and target predictability. *Quarterly Journal of Experimental Psychology, 71(*1), 64-74. https://doi.org/10.1080/17470218.2016.1247894

Veldre, A., Drieghe, D., & Andrews, S. (2017). Spelling ability selectively predicts the magnitude of disruption in unspaced text reading. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(9), 1612–1628. https://doi.org/10.1037/xhp0000425

Veldre, A., Yu, L., Andrews, S., & Reichle, E. D. (2020). Towards a complete model of reading: Simulating lexical decision, word naming, and sentence reading with Über-Reader. In Proceedings of the 42nd annual conference of the cognitive science society. Cognitive Science Society. https://hdl.handle.net/2123/22990

Venables, W. N., & Ripley, B. D. (2002). Modern Applied Statistics with S, Fourth edition. Springer, New York. ISBN 0-387-95457-0, https://www.stats.ox.ac.uk/pub/MASS4/

Vitu, F. (1991). The influence of parafoveal preprocessing and linguistic context on the optimal landing position effect. *Perception & Psychophysics*, *50*(1), 58–75. https://doi.org/10.3758/bf03212205

Vitu, F. (2011). On the role of visual and oculomotor processes in reading. *Oxford Handbooks Online*. https://doi.org/10.1093/oxfordhb/9780199539789.013.0040

Vitu, F., & McConkie, G. W. (2000). Regressive saccades and word perception in adult reading. In *Reading as a perceptual process* pp. 301-326. North-Holland.

Vitu, F., McConkie, G. W., & Zola, D. (1998). About regressive saccades in reading and their relation to word identification. In *Eye guidance in reading and scene perception* pp. 101-124. Elsevier Science Ltd.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision research, 41*(25-26), 3513-3533. https://doi.org/10.1016/S0042-6989(01)00166-3

Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & psychophysics, 47*, 583-600. https://doi.org/10.3758/BF03203111

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

Von der Malsburg, T., & Angele, B. (2017). False positives and other statistical errors in standard analyses of eye movements in reading. Journal of memory and language, 94, 119-133. https://doi.org/10.1016/j.jml.2016.10.003

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic bulletin & review, 14*, 770-775. https://doi.org/10.3758/BF03196835

Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and E–Z reader. *Cognition, 111(1),* 132–137. https://doi.org/10.1016/j.cognition.2008.12.011

Warrington, K. L., McGowan, V. A., Paterson, K. B., & White, S. J. (2018). Effects of aging, word frequency, and text stimulus quality on reading across the adult lifespan:

Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(11), 1714–1729. https://doi.org/10.1037/xlm0000543

Warrington, K.L., Xie, F., Wang, J., & Paterson, K.B. (2019). Aging effects on the visual span for alphabetic stimuli. *Experimental Aging Research, 45*, 387-399. https://doi.org/10.1080/0361073X.2019.1664465

Wen, Y., Mirault, J., & Grainger, J. (2021a). Fast syntax in the brain: Electrophysiological evidence from the rapid parallel visual presentation paradigm (RPVP). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(1), 99–112. https://doi.org/10.1037/xlm0000811

Wen, Y., Mirault, J., & Grainger, J. (2024). On relative word length and transposed-word effects. *Journal of Experimental Psychology: Human Perception and Performance, 50*(9), 934. https://doi.org/10.1037/xhp0001229

Wen, Y., Mirault, J., & Grainger, J. (2021b). The transposed-word effect revisited: The role of syntax in word position coding. *Language, Cognition and Neuroscience, 36*(5), 668–673. https://doi.org/10.1080/23273798.2021.1880608

Wen, Y., Mirault, J., & Grainger, J. (2022). A transposed-word effect on word-in- sequence identification. *Psychonomic Bulletin & Review, 29*(6), 2284–2292. https://doi.org/10.3758/s13423-022-02132-x

Wen, Y., Snell, J., & Grainger, J. (2019). Parallel, cascaded, interactive processing of words during sentence reading. *Cognition, 189*, 221–226. https://doi.org/10.1016/j.cognition.2019.04.013

Westfal, J. (2016). PANGEA: Power Analysis for General Anova designs. http://jakewestfall.org/publications/pangea.pdf

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*, 205-223. https://doi.org/10.1037/0096-1523.34.1.205

White, S. J., & Liversedge, S. P. (2006). Linguistic and nonlinguistic influences on the eyes' landing positions during reading. *Quarterly Journal of Experimental Psychology, 59*(4), 760-782. https://doi.org/10.1080/02724980543000024

White, S., & Liversedge, S. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16*(1-2), 52-78. https://doi.org/10.1080/09541440340000204

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

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*(2), 221–243. https://doi.org/10.3758/bf03196158

Winskel, H., & Perea, M. (2013). Consonant/vowel asymmetries in letter position coding during normal reading: Evidence from parafoveal previews in Thai. *Journal of Cognitive Psychology, 25*(1), 119-130. https://doi.org/10.1080/20445911.2012.753077

Winskel, H., Perea, M., & Ratitamkul, T. (2012). On the flexibility of letter position coding during lexical processing: Evidence from eye movements when reading Thai. Quarterly Journal of Experimental Psychology, 65(8), 1522-1536. https://doi.org/10.1080/17470218.2012.658409

World Literacy Foundation. (2018). The Economic & Social Cost of illiteracy. https://worldliteracyfoundation.org/wpcontent/uploads/2021/07/TheEconomicSocialCostofIlliteracy-2.pdf

Wurtz, R. H. (1996). Vision for the control of movement: the Friedenwald lecture. *Investigative Ophthalmology & Visual Science*, *37*(11), 2131-2145. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=2471263

Wurtz, R. H. (2000). Vision for the control of movement. *Cognitive neuroscience: a reader*. pp. 341-365, Blackwell.

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision research*, *48*(20), 2070-2089. https://doi.org/10.1016/j.visres.2008.03.021

Wurtz, R. H., & Goldberg, M. E. (1989). *The neurobiology of saccadic eye movements.* Elsevier.

Yan, M., Kliegl, R., Shu, H., Pan, J., & Zhou, X. (2010). Parafoveal load of word N+1 modulates preprocessing effectiveness of word N+2 in Chinese reading. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(6), 1669–1676. https://doi.org/10.1037/a0019329

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic bulletin & review*, *16*(3), 561-566. https://doi.org/10.3758/PBR.16.3.561

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sublexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(4), 1069–1075. https://doi.org/10.1037/a0026935

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2015). Perceptual span depends on font size during the reading of Chinese sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(1), 209–219. https://doi.org/10.1037/a0038097

Yan, M., Zhou, W., Shu, H., Yusupu, R., Miao, D., Krügel, A., & Kliegl, R. (2014). Eye movements guided by morphological structure: Evidence from the Uighur language. *Cognition, 132*(2), 181-215. https://doi.org/10.1016/j.cognition.2014.03.008

Yang, J. (2013). Preview effects of plausibility and character order in reading Chinese transposed words: Evidence from eye movements. Journal of Research in Reading, 36(Suppl 1), S18–S34. https://doi.org/10.1111/j.1467-9817.2013.01553.x

Yang, J., Rayner, K., Li, N., &Wang, S. (2012). Is preview benefit from word n + 2 a common effect in reading Chinese? Evidence from eye movements. *Reading and Writing, 25,* 1079–1091. https://doi.org/10.1007/s11145-010-9282-7

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

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from word n + 2? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1192–1204. https://doi.org/10.1037/a0013554

Yatabe, K., Pickering, M. J., & McDonald, S. A. (2009). Lexical processing during saccades in text comprehension. *Psychonomic bulletin & review, 16*, 62-66. https://doi.org/10.3758/PBR.16.1.62

Yoshida, K., Iwamoto, Y., Chimoto, S., & Shimazu, H. (1999). Saccade-related inhibitory input to pontine omnipause neurons: an intracellular study in alert cats. *Journal of neurophysiology*, *82*(3), 1198-1208. https://doi.org/10.1152/jn.1999.82.3.1198

Zang, C. (2019). New perspectives on serialism and parallelism in oculomotor control during reading: The multi-constituent unit hypothesis. *Vision, 3*(4), 50. https://doi.org/10.3390/vision3040050

Zang, C., Zhang, Z., Zhang, M., Degno, F., & Liversedge, S. P. (2023). Examining semantic parafoveal-on-foveal effects using a Stroop boundary paradigm. *Journal of Memory and Language, 128*, 104387. https://doi.org/10.1016/j.jml.2022.104387

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. https://doi.org/10.1037/xhp0000644

Appendix A

Supplementary table 1.

Descriptive statistics on Frequency for all positions in the full set of 108 sentences (Experiment 1 – Chapter 2). Frequency data was obtained and calculated via the Log Zipf scale (van Heuven et al., 2014).

Word	Frequency		t-Test Results for the Critical Comparison between Words				
	Range	M (SD)	First Word t-value (p-value)	Pre-Target t-value (p-value)	Target Word 1 t-value (p-value)	Target Word 2 t-value (p-value)	Post-Target Grammatical t-value (p-value)
First Word	2.11 - 7.67	5.76 (1.43)					
Pre-Target	2.95 - 6.55	5.34 (0.68)	2.79 (0.006)				
Target Word 1	3.40 - 6.90	5.49 (0.71)	1.79 (0.075)	-1.59 (0.114)			
Target Word 2	3.51 - 7.19	5.50 (0.82)	1.66 (0.098)	-1.58 (0.116)	0.11 (0.911)		
Post-Target Grammatical	2.37 - 6.45	4.81 (0.81)			6.57 (<0.001)	6.24 (<0.001)	
Post-Target Ungrammatical	2.39 - 6.35	4.63 (0.74)			8.79 (<0.001)	8.28 (<0.001)	-1.78 (0.076)

Supplementary table 2.

T-tests on pre-screens for the full set of 108 stimuli (Experiment 1 – Chapter 2)..

Test/ Pre-screen	Naturalness	
	t	*p*
Grammatical and Transposed vs Grammatical and Not Transposed	-56.88	< 0.001
Ungrammatical and Transposed vs Ungrammatical and Not Transposed	-4.71	< 0.001
Grammatical and Transposed vs Ungrammatical and Transposed	7.32	< 0.001
Grammatical and Not Transposed vs Ungrammatical and Not Transposed	64.28	< 0.001
	Predictability	
Target Word 1 vs. Target Word 2	-1.12	0.266

Supplementary table 3.

Descriptive Statistics for all measures based on the full set of 108 stimuli (Experiment 1 – Chapter 2).

Measure/ Condition	Grammatical and Not Transposed	Grammatical and Transposed	Ungrammatical and Not Transposed	Ungrammatical and Transposed
Accuracy	93 (26)	92 (28)	92 (28)	99 (11)
RTs	2132 (773)	2279 (837)	2249 (805)	2160 (780)
	First Word			
	Not Transposed		Transposed	
FFD	182 (55)		180 (53)	
SFD	178 (53)		177 (52)	
GD	202 (93)		207 (101)	
RP	11 (32)		13 (34)	
	Grammatical and Not Transposed	Grammatical and Transposed	Ungrammatical and Not Transposed	Ungrammatical and Transposed
TVT	241(130)	232 (129)	225 (130)	224 (122)
RIn	25 (43)	15 (36)	16 (36)	16 (36)
	Pre-Target			
	Not Transposed		Transposed	
FFD	197 (52)		194 (52)	
SFD	203 (52)		201 (53)	
GD	239 (88)		237 (89)	
SP	8 (26)		8 (28)	
RP	23 (42)		23 (42)	
ROut	4 (20)		5 (21)	
GPT	250 (102)		249 (105)	
	Grammatical and Not Transposed	Grammatical and Transposed	Ungrammatical and Not Transposed	Ungrammatical and Transposed
TVT	352 (187)	348 (203)	321 (189)	322 (194)
RIn	42 (49)	37 (48)	25 (43)	28 (45)
	Target Word 1			
FFD	219 (63)	222 (69)	217 (60)	221 (68)
SFD	230 (60)	245 (78)	229 (65)	241 (78)
GD	254 (89)	275 (111)	257 (93)	270 (109)
SP	13 (34)	13 (33)	14 (32)	12 (32)
RP	17 (38)	21 (38)	18 (39)	21 (41)
ROut	6 (24)	8 (27)	5 (21)	9 (28)
GPT	275 (120)	308 (153)	271 (120)	297 (146)
TVT	397 (231)	466 (258)	371 (221)	436 (245)
RIn	44 (50)	51 (50)	36 (48)	45 (50)
	Target Word 2			
FFD	250 (83)	253 (85)	251 (84)	252 (91)
SFD	282 (95)	301 (100)	288 (97)	295 (100)
GD	328 (143)	345 (170)	330 (146)	328 (153)
SP	9 (28)	6 (24)	9 (29)	7 (25)
RP	29 (45)	31 (46)	28 (45)	28 (45)
ROut	24 (43)	34 (48)	13 (33)	29 (45)
GPT	430 (273)	539 (318)	390 (221)	471 (266)
TVT	465 (235)	543 (266)	515 (253)	503 (254)
RIn	30 (46)	26 (44)	46 (50)	30 (46)
	Post-Target			
FFD	23 (94)	242 (105)	(257 (95)	238 (94)
SFD	249 (100)	245 (112)	273 (105)	250 (111)
GD	309 (160)	306 (175)	366 (188)	304 (164)
FP	79 (41)	72 (45)	90 (30)	84 (37)
RP	26 (44)	23 (42)	37 (48)	23 (42)
ROut	68 (47)	63 (48)	70 (46)	56 (50)
GPT	758 (570)	789 (641)	841 (594)	675 (550)
TVT	475 (235)	369 (243)	453 (241)	369 (218)

Note: The values for FFD; SFD; GD; GPT, TVT and RT are given in milliseconds, while the values for Accuracy, SP; RP; FP; RIn and ROut are given in percentages.

Supplementary table 4.

Fixed effects estimates from the generalised mixed effects models based on the full set of 108 stimuli (Experiment 1 – Chapter 2).

Measure/ Condition	Intercept				Transposition (Transposed-Not Transposed)				Final Word Grammaticality (Grammatical-Ungrammatical)				Interaction			
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
Accuracy	3.93	0.22	**18.09**	[3.51, 4.36]	1.24	0.39	**3.16**	[0.47, 2.01]	-1.61	0.45	**-3.59**	[-2.50, -0.73]	-2.92	0.82	**-3.57**	[-4.52, -1.31]
RTs	2405.55	5.84	**411.88**	[2394.10,2416.99]	41.30	7.62	**5.42**	[26.36,56.24]	12.81	6.52	*1.96*	[0.02,25.59]	264.39	5.55	**47.68**	[253.52,275.24]
								First Word								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z) -value*	*Conf Int*	*β*	*SE*	*t(z) -value*	*Conf Int*	*β*	*SE*	*t(z) -value*	*Conf Int*
FFD	182.90	6.18	**29.58**	[170.79,195.02]	-1.41	1.18	-1.20	[-3.72,0.90]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	192.93	8.55	**22.56**	[176.17,209.68]	-2.40	2.16	-1.11	[-6.63,1.84]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	207.00	9.22	**22.44**	[188.93,225.08]	3.39	2.90	1.17	[-2.29,9.07]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
RP	-3.35	0.27	**-12.25**	[-3.89, -2.82]	0.06	0.13	0.44	[-0.20,0.31]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	235.13	6.44	**36.54**	[222.52,247.74]	-7.01	4.24	-1.65	[-15.32,1.30]	10.77	4.87	2.21	[1.22,20.33]	-9.50	6.01	-1.58	[-21.28,2.28]
RIn	-1.85	0.19	**-10.02**	[-2.22, -1.49]	-0.32	0.10	**-3.10**	[-0.52, -0.12]	0.43	0.13	**3.27**	[0.17,0.69]	-0.69	0.21	**-3.40**	[-1.09, -0.28]
								Pre-Target								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	200.23	4.71	**42.49**	[190.99,209.47]	-3.11	2.37	-1.31	[-7.74,1.53]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SFD	210.24	6.47	**32.48**	[197.56,222.93]	-2.13	3.53	-0.60	[-9.04,4.78]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GD	242.06	7.35	**32.94**	[227.66,256.47]	-2.17	3.61	-0.60	[-9.24,4.90]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SP	-3.31	0.25	**-13.47**	[-3.79, -2.83]	0.10	0.13	0.75	[-0.15,0.34]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
RP	-1.50	0.17	**-8.80**	[-1.84, -1.17]	-0.01	0.08	-0.06	[-0.17,0.16]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ROut	-3.61	0.19	**-18.82**	[-3.98, -3.23]	0.08	0.16	0.47	[-0.24,0.40]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
GPT	256.85	8.62	**29.81**	[239.96,273.74]	-1.51	4.55	-0.33	[-10.43,7.41]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
TVT	349.45	9.83	**35.57**	[330.19,368.71]	3.10	6.36	0.49	[-9.35,15.56]	34.09	6.44	**5.30**	[21.47,46.71]	2.98	6.63	0.45	[-10.01,15.96]
RIn	-0.95	0.20	**-4.81**	[-1.34, -0.57]	0.03	0.11	0.22	[-0.20,0.25]	0.78	0.11	**7.33**	[0.57,0.99]	-0.53	0.16	**-3.33**	[-0.85, -0.22]
								Target Word 1								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	223.48	4.63	**48.32**	[2114.42,232.55]	4.27	3.63	1.18	[-2.84,11.38]	1.18	2.90	0.41	[-4.49,6.86]	-1.74	4.87	-0.36	[-11.28,7.80]
SFD	240.09	6.85	**35.05**	[226.67,253.52]	14.39	6.31	*2.28*	[2.02,26.76]	4.14	5.51	0.75	[-6.66,14.94]	0.39	10.26	0.04	[-19.72,20.50]
GD	265.51	5.74	**46.28**	[254.27,276.76]	18.10	5.62	**3.21**	[7.09,29.11]	-0.09	4.60	-0.02	[-9.11,8.93]	4.47	6.79	0.66	[-8.84,17.79]
SP	-2.61	0.23	**-11.46**	[-3.06. -2.16]	-0.06	0.10	-0.53	[-0.26,0.15]	0.06	0.10	0.55	[-0.14,0.26]	0.17	0.20	0.86	[-0.22,0.57]
RP	-1.70	0.14	**-11.85**	[-1.99, -1.42]	0.26	0.09	**2.99**	[0.09,0.44]	0.01	0.09	-0.07	[-0.16,0.18]	0.08	0.17	0.45	[-0.26,0.42]
ROut	-2.96	0.15	**-19.85**	[-3.25, -2.67]	0.51	0.13	**3.91**	[0.26,0.77]	0.13	0.13	1.00	[-0.13,0.39]	-0.37	0.26	-1.42	[-0.89,0.14]
GPT	291.83	5.81	**50.27**	[280.45,303.21]	33.90	5.55	**6.11**	[23.02,44.78]	8.96	4.73	1.89	[-0.31,18.24]	9.32	5.71	1.63	[-1.87,20.51]
TVT	429.96	20.26	**21.22**	[390.24,469.68]	74.86	6.15	**12.17**	[62.80,86.92]	31.14	5.88	**5.30**	[19.62,42.66]	8.89	6.56	1.36	[-3.96,21.74]
RIn	-0.33	0.15	-2.19	[-0.62, -0.03]	0.42	0.09	**4.85**	[0.25,0.59]	0.38	0.08	**4.65**	[0.22,0.54]	-0.16	0.15	-1.09	[-0.44,0.13]
								Target Word 2								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	255.74	5.83	**43.87**	[244.32,267.17]	2.48	4.45	0.56	[-6.23,11.20]	-0.69	3.64	-0.19	[-7.83,6.45]	-3.34	7.16	-0.47	[-17.96,10.69]
SFD	293.04	8.52	**34.40**	[276.34,309.73]	14.97	8.26	1.81	[-1.21,31.16]	1.09	9.55	0.11	[-17.62,19.81]	10.13	9.40	1.08	[-8.31,28.56]
GD	337.03	7.30	**46.20**	[322.73,351.32]	12.13	6.43	1.89	[-0.47,24.74]	7.81	6.37	1.23	[-4.67,20.29]	15.07	7.49	*2.01*	[0.39,29.75]
SP	-3.03	0.20	**-15.34**	[-3.41, -2.64]	-0.34	0.17	-2.03	[-0.67,0.01]	-0.06	0.17	-0.34	[-0.38,0.27]	-0.09	0.34	-0.25	[-0.76,0.58]
RP	-0.99	0.11	**-9.35**	[-1.19, -0.78]	0.06	0.07	0.86	[-0.08,0.20]	0.08	0.08	1.04	[-0.07,0.23]	0.18	0.14	1.30	[-0.09,0.46]
ROut	-1.34	0.13	**-10.18**	[-1.59, -1.08]	0.87	0.11	**8.27**	[0.66,1.08]	0.61	0.08	**7.35**	[0.45,0.77]	-0.60	0.17	**-3.63**	[-0.92, -0.28]
GPT	467.20	6.90	**67.74**	[453.68,480.72]	103.33	9.97	**10.37**	[83.79,122.87]	64.10	11.02	**5.82**	[42.51,85.69]	33.13	6.45	**5.14**	[20.49,45.78]
TVT	519.84	36.92	**14.08**	[447.48,592.20]	35.05	8.14	**4.30**	[19.09,51.00]	-7.33	7.19	-1.02	[-21.43,6.77]	95.69	9.67	**9.90**	[76.74,114.65]

RIn	-0.88	0.14	**-6.22**	[-1.16, -0.61]	-0.52	0.07	**-6.93**	[-0.66, -0.37]	-0.50	0.08	**-6.04**	[-0.66, -0.34]	0.61	0.15	**4.10**	[0.32,0.90]
								Post-Target								
	β	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*	*β*	*SE*	*t(z)-value*	*Conf Int*
FFD	245.09	6.64	**36.91**	[232.08,258.11]	-13.59	5.17	***-2.63***	[-23.73, -3.45]	-7.93	5.08	-1.56	[-17.88,2.03]	20.95	6.34	**3.30**	[8.52,33.38]
SFD	269.57	10.42	**25.87**	[249.16,290.01]	-20.32	8.26	*-2.46*	[-36.52, -4.12]	-20.71	9.76	-2.12	[-39.84, -1.58]	18.59	11.59	*1.60*	[-4.12, 41.31]
GD	318.31	6.74	**47.24**	[305.10,331.51]	-40.78	4.92	**-8.29**	[-50.42, -31.14]	-37.38	5.63	**-6.64**	[-48.41, -26.35]	65.32	5.48	**11.92**	[54.57,76.06]
FP	2.04	0.20	**10.07**	[1.64,2.44]	-0.48	0.13	**-3.75**	[-0.73, -0.23]	-1.02	0.13	**-7.78**	[-1.28, -0.77]	0.30	0.19	1.53	[-0.08,0.68]
RP	-1.33	0.16	**-8.58**	[-1.63, -1.02]	-0.55	0.12	**-4.77**	[-0.78, -0.33]	-0.45	0.12	**-3.61**	[-0.69, -0.20]	0.57	0.22	**2.56**	[0.14,1.01]
ROut	0.73	0.17	**4.28**	[0.39,1.06]	-0.52	0.10	**5.14**	[-0.72, -0.32]	0.09	0.12	0.76	[-0.14,0.31]	0.38	0.17	2.28	[0.05,0.70]
GPT	745.87	7.23	**103.23**	[731.70,760.03]	-86.82	8.33	**-10.42**	[-103.14, -70.50]	-5.20	6.82	-0.76	[-18.57,8.17]	190.71	11.33	**16.83**	[168.49,212.92]
TVT	386.03	7.87	**49.06**	[370.61,401.45]	-53.47	6.20	**-8.63**	[-65.61, -41.33]	-55.01	6.47	**-8.51**	[-67.69, -42.34]	88.96	8.73	**10.19**	[71.85,106.08]

Note. Significant terms are presented in bold, and terms approaching significance are underlined. All *p* values were adjusted via the Bonferroni correction with the summary(glht(<modelName>), test = adjusted(“bonferoni”)) function. Additionally, terms which were qualitatively different between the analyses with the full set and the reduced set of stimuli, are presented in italics. N/A signifies that the factor or interaction was not considered in the model.

^Confidence intervals are calculated with the *confint* function in R with method = *Wald* at the 2.5 percentile and 97.5 percentile.