UNIVERSITY OF SOUTHAMPTON

FACULTY OF SOCIAL, HUMAN, AND MATHEMATICAL SCIENCES

Psychology

Oculomotor Control during Reading

by

Michael G. Cutter

Thesis for the degree of Doctor of Philosophy

July 2016

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL, HUMAN, AND MATHEMATICAL SCIENCES

Psychology

Thesis for the degree of Doctor of Philosophy

Oculomotor Control during Reading

Michael G. Cutter

The work presented in this thesis explores several issues relevant to the understanding of oculomotor control during reading, and theoretical accounts of this issue. In Chapter 1 I give a summary of prior research investigating the factors that determine where and for how long the eyes fixate during reading, and how information is integrated across multiple fixations. Furthermore, I outline how current models of oculomotor control during reading can account for these various findings. The research I present in the following empirical chapters then aims to extend our understanding of oculomotor control during reading. In Chapter 2 I present an investigation into whether spaced compound words, such as *teddy bear*, are processed in a similar manner to single long words or as two separate words. It is found that these multi-word units are indeed processed as single long words, with information being processed earlier than would be expected of two separate words. In Chapter 3 I investigate whether orthographic information from one parafoveal word influences the processing of an adjacent parafoveal word, finding that the processing of information from these two words seems to be independent. In Chapter 4 I examine the extent to which the oculomotor targeting system is flexible in relation to the mean word length in the text currently being read, finding that both word skipping behaviour and the saccade length at which readers would most accurately land in the centre of an upcoming word adapted. Finally, in Chapter 5 I explore how well the models of oculomotor control outlined in Chapter 1 are able to account for my findings, as well as discussing the common implications of these studies for our theoretical understanding. Finally, I outline some ways in which the research presented can be extended in the future to further increase our knowledge of oculomotor control during reading.

Contents

List of Tables

List of Figures

Declaration of Authorship

I, *Michael G. Cutter,* declare that the thesis entitled *Oculomotor Control during Reading* and the work presented in it are my own and has been generated by me as the result of my own original research. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Either none of this work has been published before submission, or parts of this work have been published as:

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced compounds. *Journal of Experimental Psychology: Learning Memory and Cognition, 40,* 1778-1786. doi:10.1037/xlm0000013

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across fixations in reading? In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 245-260). New York, NY: Oxford University Press.

Signed: ……………………………………………………………………..

Date:……………………………………………………………………………..

Acknowledgements

First and foremost, I would like to thank my supervisors, Simon Liversedge and Denis Drieghe. Without their help over the past five years none of the work presented in this thesis would have been possible. They have both always been happy to spend time helping me, whether it be with theoretical, technical, or statistical issues. They have also both spent a great deal of time feeding back on my writing. I hope that looking back over my work they will consider the time spent helping me to have been worthwhile.

The majority of work in this thesis has been conducted with a mind to publication, with this having a large influence on my choice to submit this as a three-paper thesis. Parts of the introduction have been published, alongside all of Chapter 2. Chapter 4 has already been through one round of reviews, and it is the revised version that appears in this thesis. I have also presented the work forming all of the empirical chapters at conferences at least once. Going through the review process, as well as conversations with other researchers at conferences, has been a vital part in the development of my ideas, and so I would like to thank anybody who has provided feedback on my work, including reviewers, editors, and fellow conferences attendees.

I would also like to thank the various members of the Centre for Vision and Cognition, the Psychology department at the University of Southampton in general, and various people who I have shared an office with over the last five years. Too many people to list have been a great deal of help both academically and personally. This also includes members of the eye-tracking group at Tianjin Normal University.

I would also like to thank the University of Southampton’s Mayflower Scholarship scheme, for having funded my candidature.

Finally, I would like to thank various family members and friends who have provided a pleasant distraction and forms of fun at times when the stress was getting to be just a little too much.

Abbreviations

FFD- First fixation duration

GD- Gaze duration

GP- Go-past time

LME- Linear mixed effect

LMM- Linear mixed model

LP- Landing position

MWU- Multi-word unit

PWF- Proportion of words fixated

SFD- Single fixation duration

U3- Uniform sentences of three letter words

U4- Uniform sentences of four letter words

U5- Uniform sentences of five letter words

NU- Non-uniform sentences

Chapter 1

Introduction

Large portions of the following literature review have been published as Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across fixations in reading? In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 245-260). New York, NY: Oxford University Press. The majority of this material forms section 1.2. Supplementary material has been added which relates to material which was not relevant to this paper but will be relevant to this thesis.

1.1.1. The Necessity of Eye Movements during Reading

The anatomy of the human eye and the retina in particular places a large level of constraint on the manner in which people are able to sample any visual stimulus. This includes any text that a person is trying to read and comprehend. The surface of the retina is made up of two different types of photoreceptors. The first type is cones, and it is these that allow for the perception of visual details in the form of high spatial densities; crucially, this includes the features of the letters that form words. On the other hand, rods primarily provide coarser visual information, and are not typically adequate for obtaining detailed visual information. These two types of photoreceptor are not equally distributed across the retina, and as such certain parts of the human visual field are more suited to sampling detailed visual information than others. Specifically, there is a 2 degree area of the retina in which the cones are most densely packed, known as the fovea. We are able to extract the most detailed, fine grained visual information from stimuli which project light onto the fovea. Beyond the fovea the density of cones and thus visual acuity is considerably reduced, in the parafovea. The parafovea extends a further 4 degrees of visual angle to either side of the fovea (Balota & Rayner, 1991). The drop in visual acuity continues beyond the parafovea, into the periphery.

Due to the constraints imposed by the anatomy of the retina, it is necessary to take multiple spatially distributed samples of a body of text. During reading the eye will make what are called saccadic eye movements, which serve to move the fovea from one location in the text to another. As will become clear in the following sections, this is done in a highly systematic manner in order to ensure that the eye provides the cognitive processing system with the visual information necessary to successfully read, in a time course ideal for the processing difficulty of the text being read. The periods of stillness in between saccades are referred to as fixations, and it is during these fixations that visual information is extracted from the page.

1.1.2. Eye-Trackers and their Capabilities

It is possible to track the pattern of fixations and saccades that people make across text with a high degree of temporal and spatial accuracy using an eye-tracker. An eye-tracker will sample the position of the eye a large number of times during a single second. In the specific case of the SR Research Eyelink 1000, which was used for data acquisition in the research presented in this thesis, 1000 samples are taken per second. This sample will typically be spatially accurate to about half a character, assuming proper calibration. Through a complex series of filters and algorithms the eye-tracker's software reduces and segments this constant stream of samples into a more interpretable set of fixations and saccades. At a very basic level, the software associated with the tracker uses an algorithm to identify saccades as periods during which the measured position of the eye is changing enough from sample to sample to be indicative of the eye being in transit. The samples that occur in between these saccades are then grouped together as part of a fixation. Further operations are performed on these data to calculate a large number of variables describing the spatial and temporal patterns of fixations and saccades.

Due to the accuracy with which eye-trackers allow us to measure the position of the eye, it is possible to implement what are known as gaze-contingent paradigms. Research making use of these paradigms will be discussed later, and they are also used extensively in the original work presented in this thesis. Essentially, the eye-tracker uses the current measured position of the eye in order to alter the visual information which is available from the display monitor. Typically, the display change will occur prior to the completion of a saccade, meaning that participants are not consciously aware of the manipulation, due to visual information not being encoded during a saccade (see Martin, 1974).

```
            *
Lbo dhr quickly jumped over lfa pameo ez bg pem tvoq
                    *
Lbo dhr pvlsziv jumped over the feneo ez bg pem tvog
                        *
Lbo dhr pvlsziv fbtyed over the fence az bg pem tvog
```

Figure 1.1. An illustration of the moving window paradigm. The point of fixation is represented by the asterisk. In the current illustration four characters to the left and fourteen characters to the right of the point of fixation are available.

There are several ways in which gaze-contingent techniques can be used. For example, one way in which this technique has been used is to implement the moving window paradigm (McConkie & Rayner, 1975). In this paradigm, a window of normal text is set by the experimenter around the point of fixation. Within this window the characteristics of the text being read are preserved, whereas outside the window the text is masked. As a saccade is made a display change occurs, so that a new window of unmasked text is set around the new point of fixation (see Figure 1.1). The smallest window size for which reading occurs at a rate similar to normal reading is referred to as the perceptual span. As such, this paradigm gives an estimate of how much information readers extract during a single fixation, albeit with no indication about the form of the extracted information. For English, this extends 3-4 character spaces to the left and 14-15 characters to the right of fixation (McConkie & Rayner, 1975; McConkie & Rayner, 1976).

A second type of gaze-contingent technique is the boundary paradigm (Rayner, 1975). In the boundary paradigm an invisible boundary is set (usually) at the end of a pre-target word (see Figure 1.2). Prior to the eyes crossing the boundary, a preview string is presented instead of the target word. This preview can be the target word itself (an identity preview), a different word, or a non-word. The preview quickly changes to the target word as a saccade is made that crosses the boundary. As will be seen in Section 1.2, this technique has been widely used and has demonstrated that when readers are given an identity preview of a target word, they take less time to process and identify it relative to when they are given an incorrect preview. This advantage is referred to as the *preview benefit*. The fact that readers gain a preview benefit strongly suggests that they have extracted and processed information about the preview string before fixating it, and then integrated this information with information obtained on the next fixation, usually made on the word itself.

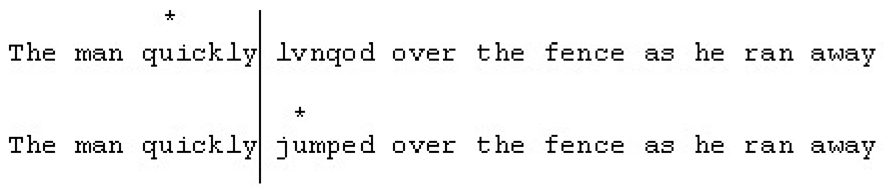


Figure 1.2. An illustration of the boundary paradigm. The point of fixation is represented by the asterisk. The invisible boundary is represented by the black line. Prior to the point of fixation crossing this boundary there is a non-word preview of the upcoming word. As the point of fixation crosses the boundary, the target word replaces the preview.

The two gaze-contingent techniques outlined above both involve manipulating the availability of information outside of the fovea, and will be discussed extensively in Section 1.2. However, for now I will focus upon the importance of information being available foveally. A third gaze-contingent technique was used by Rayner and Bertera (1979) to demonstrate the importance of foveal processing during reading. In this study, a gaze-contingent paradigm was used to replace the text at and around the point of fixation with an x-string mask of varying sizes. Thus, the extent to which information was available to process in the fovea was varied. With no mask participants read 332 words per minute, whereas when a single letter in the fixated word was masked this dropped to 165 words per minute. This drop-off continued, with masks of three and five characters leading to reading rates of 55 and 42 words per minute, respectively. Furthermore, when asked to report what they had read participants often reported incorrect versions of the sentence. Thus, the findings of this study suggest that in order for words to be correctly identified it is vital for them to be processed in foveal vision. If this is not possible words are identified far more slowly, and often incorrectly. It should be noted that some words can be identified without being directly fixated; this will be discussed in Section 1.1.4.

1.1.3. Determinants of Fixation Durations during Reading

One experimental finding that demonstrates the sensitivity of eye movements to the processing difficulty of a word is the frequency effect (Altarriba, Kroll, Sholl, & Rayner, 1996; Henderson & Ferreira, 1990; Just & Carpenter, 1980; Liversedge, Rayner, White, Vergilino-Perez, Findlay, and Kentridge, 2004; Rayner & Fischer, 1996; Reingold, Reichle, Glaholt, & Sheridan, 2012; Schilling, Rayner, & Chumbley, 1998; White, 2008; see

Rayner, 1998 for an extensive review). Before proceeding it is important to outline why exactly it is that a word's frequency would influence how difficult it is to identify, and thus why we would expect to see an effect of this variable on reading times. While there are many ways to address this issue I will focus upon an argument advanced by Reichle and Perfetti (2003). These scholars attempt to explain word identification within the more general framework of episodic memory. Essentially, each time a word is encountered a memory trace will be created, which, among other information, includes both the perceptual and semantic characteristics of the word. A memory trace will be created in long-term memory each time a word is encountered, and due to this a reader will acquire a large number of rich representations of the words making up their language. These memory traces can then be used to aid word recognition, and the recall of the various characteristics associated with a word. Words that are encountered more frequently will have richer and more numerous traces in long-term memory, and as such can be recognised more quickly during future encounters than words that are encountered less frequently. In support of this argument, Reichle and Perfetti (2003) demonstrated that the ability of an existing model of memory (MINERVA 2; Hintzman, 1984) to identify and retrieve the characteristics of a set of words was strongly influenced by these words' frequencies. Thus, one reason why frequency would be expected to influence how long a word is fixated for is that this variable influences the ease with which a word can be retrieved from long term memory.

It has repeatedly been found that how frequently a word occurs in a language influences how long participants fixate this word for, with low frequency words being fixated for longer than high frequency words. This has been found over many studies, using both controlled experimental designs, and less controlled corpus-based analyses. At this point it is worth briefly outlining the distinction between these two different approaches. In the case of a controlled experimental design, the effect of frequency on fixation times has been tested by varying a single target word in a sentence on the basis of its frequency, while controlling for other word-based factors across experimental conditions. Participants are typically shown a number of sentences including either a high or low frequency target word, with twelve items in each condition being a fairly standard number. Fixation durations on the target words are then examined as a function of whether they are of high or low frequency. As an example, in one study examining frequency effects participants were presented with the sentence *Mary bought a chest/trunk despite the high price* with *chest* being the high frequency target, and *trunk* being the low frequency target. On average, participants remained fixated on the high frequency word for less time than the low frequency word (Henderson & Ferreira, 1990). Due to the controlled nature of experimental designs, it is generally easy to be sure that an effect was caused by the variable that was specifically manipulated.

Similar effects have also been examined in larger scale corpus studies (Kliegl, Grabner, Rolfs, & Engbert, 2004). In these corpus studies there are no target words per se. Rather, participants will read a large amount of text which has not been manipulated in any specific way. Each word in this corpus will vary on a large number of dimensions typical of natural language, with words freely varying on factors such as frequency, length, predictability etc. When this data is analysed, the characteristics of each word can be entered into the analysis as predictor variables, and it is possible to determine which of these factors influences fixation times in a significant way. One major advantage of this approach is that it allows researchers to address several questions using a single data set, by examining fixation times as a function of the many different factors that have been allowed to vary freely. Furthermore, it typically provides a greater deal of statistical power, due to every word in the sentence being analysed, rather than a single target word in each sentence. However, as will be seen in Section 1.2.7 there are also some major drawbacks to this approach. For now I will defer discussing these drawbacks, until they become relevant to the interpretation of a corpus-based finding. To return to frequency effects, Kliegl et al. (2004) did indeed observe frequency effects, with fixation times increasing as word frequency decreased.

Remarkably, these frequency effects have been found to endure in experiments using the disappearing text paradigm (Liversedge, Rayner, White, Vergilino-Perez, Findlay, & Kentridge, 2004). In this study each word in the text disappeared 60ms after participants first fixated them, and did not re-appear until after a saccade was made to a different word. Even under these unusual conditions the amount of time participants remained fixated upon a target word was still heavily influenced by the word’s frequency, despite the word being absent during the majority of that fixation. This finding reinforces the notion that there is a tight link between the difficulty of lexically processing a word and how long people fixate that word. A further cognitive factor that influences eye movements is the predictability of a word from the preceding sentential context. The predictability of a target word in a sentence is typically assessed by presenting a separate group of participants with the target sentence up until the pre-target word, and asking them to predict the next word in the sentence. If many participants predict a word it is considered highly predictable, whereas if few predict a certain word it is considered unpredictable. Rayner and Well (1996) investigated this effect by presenting participants with a sentence which included either a highly predictable or unpredictable target word. For example, in the sentence *the hikers slowly climbed up the mountain/hillside to get a better view*, *mountain* is highly predictable and *hillside* is not. Participants' eye movements were measured as they read these sentences. Participants fixated the predictable words for less time than they fixated the unpredictable words. This further demonstrates that the amount of time participants fixate a word is influenced by its processing difficulty, and that this includes the processing difficulty of a word in the context of a particular sentence.

There are many additional factors which influence fixation durations during reading. One such variable is age of acquisition (Juhasz & Rayner, 2003; Juhasz & Rayner, 2006), which is the age at which a reader first encounters a word. Another is the rated familiarity of a word (Williams & Morris, 2004), which is a measure of how commonly a reader believes they personally encounter a word. Word length has also been found to influence fixation times upon a word, with longer words being fixated for longer than shorter ones (Hautala, Hyönä, & Aro, 2011; McDonald, 2006; Rayner, Sereno, & Raney, 1996). The plausibility of a word in the sentential context also has an effect upon how long it is fixated for (Rayner, Warren, Juhasz, & Liversedge, 2004). In addition to these variables there are a number of orthographic factors which influence word identification times. These include the orthographic familiarity of a word, which is a measure of how commonly the letters within that word tend to appear together (White, 2008), and the size of a word's orthographic neighbourhood, which is the number of other words that can be made from one word by changing a single letter (Pollatsek, Perea, Binder, 1999; see Perea, 2015, for a wider definition of neighbourhood size).

From the above research it should be clear that factors related to the difficulty of processing a word heavily influence the amount of time it is fixated, with words that are difficult to process remaining fixated for longer. Thus, the temporal characteristics of eye movements are clearly tightly linked to on-going cognitive processing when reading.

1.1.4. Determinants of Fixation Locations during Reading

While the above studies demonstrate that fixation durations are influenced by cognitive processes, they do not show whether where the eyes fixate during reading is determined by similar factors. It could be the case that where the eyes fixate is influenced by a similar range of high level cognitive influences as how long a word is fixated for (e.g. frequency, predictability, familiarity, etc.). Alternatively, it could be that saccadic targeting is more strongly influenced by low-level, visual factors, with higher level word characteristics not influencing eye-movement behaviour until direct fixation. By low-level characteristics I refer to the influence of largely visual factors, such as word length, visually salient letter combinations, or even the characteristics of the oculomotor control system determining fixation locations. Broadly speaking, there are two levels at which saccadic targeting can be examined. The first is which words in a text are fixated, and the second is the location of a fixation within a word.

Approximately one third of words are skipped during reading (Brysbaert, Drieghe, & Vitu, 2005). Word skipping is highly systematic, such that there are a number of factors which reliably influence whether or not a word is directly fixated. One of the strongest predictors of skipping is word length. For example, Vitu, O' Regan, Inhoff, and Topolski (1995) found that in their study 60% of three letter words were skipped, whereas only 30% of five letter words were skipped. In a further study Hautala, Hyönä, and Aro (2011) found that this effect is driven more by the spatial extent of the word, rather than the number of letters. They investigated this by presenting a four or six letter target word in a sentence, while varying the font in such a way that the words of different length would either subtend the same amount of visual angle, or differing amounts of visual angle. They found that participants would skip the six letter word less when it subtended a greater amount of visual angle than the four letter word, but not when it subtended the same amount of visual angle. The distance of the target word from the current fixation location also has a large influence on word skipping rates. For example, in one study a five letter word was skipped less than 10% of the time when the prior fixation was seven characters from its start, but around 25% of the time when the prior fixation was only three characters away from the start of the word (Rayner, Sereno, & Raney, 1996). It should be clear that the influence of both word length and launch site constitute fairly large effects, with a 30% change in skipping due to a two letter change in length, and a 15% change due to a four letter change in launch site.

In comparison, higher level cognitive factors such as frequency and predictability have relatively modest effects on skipping. In a meta-analysis, Brysbaert et al. (2005) found that on average a high frequency word is skipped 16% of the time, and a low frequency word 11% of the time. This effect depends upon the length of the word, with a 10% difference between high and low frequency five letter words, but a difference of only

4% for seven letter words (Rayner & Fischer, 1996). Brysbaert et al. also found that on average participants will skip a predictable word 8% more often than an unpredictable word. This effect of predictability on skipping is rapid, with Fitzsimmons and Drieghe (2013) observing that a word which only becomes predictable given the preceding word was skipped as often as one which gradually becomes predictable across an entire sentence. A final factor that affects word skipping is syllabic structure, with monosyllabic words (e.g. *phone*) being skipped 6% more than matched disyllabic (e.g. *music*) words (Fitzsimmons & Drieghe, 2011).

Where a reader fixates within a word is solely driven by low level visual and oculomotor factors, at least in an alphabetic language such as English (see Yan, Zhou, Shu, Yusupu, Miao, Krügel, & Kliegl, 2014 for a demonstration of morphological structure influencing saccadic targeting in Uighur). Rayner (1979) observed that fixation positions within a word followed a normal distribution, with the most common landing position being slightly to the left of the word's centre. This modal landing position is referred to as the *preferred viewing location*. One of the primary drivers of the preferred viewing location is the distance between the launch site of the incoming saccade, and the centre of the word being targeted. McConkie, Kerr, Reddix, and Zola (1988) found that when a saccade was launched from less than seven characters away readers tended to overshoot the centre of a word, whereas a saccade launched from further away tended to undershoot the centre of a word. In other words, the further away from the fixation target a saccade is launched, the further the preferred viewing location will shift to the left portion of a word. These effects are due to a saccadic range error, whereby readers have a bias for making saccades of a certain length (the *preferred saccade length*). In English, this is seven characters. For each character that the saccade launch site is further or nearer from the centre of the upcoming word than the preferred saccade length, there will be approximately half a character of undershoot or overshoot. Furthermore, longer saccades also tend to lead to a greater spread of landing positions within a word, due to increased motor error for these long movements. Thus, McConkie et al.'s work showed that one of the primary determinants of where the eye lands within a word is the location of the previous fixation.

The visual features of an upcoming word have also been shown to affect landing positions within words. A number of studies (Plummer & Rayner, 2012; Radach, Inhoff, & Heller, 2004; White & Liversedge, 2004; White & Liversedge, 2006a; White & Liversedge 2006b) have found that when the beginning of an upcoming word features unusual orthographic sequences (e.g. *ngricultural*) initial fixation locations are closer to the start of a word than when this is not the case (e.g. *agricultural*). While this low level orthographic information affects saccadic targeting, higher level factors such as lexical status (Plummer & Rayner, 2012) or word frequency (Rayner, Sereno, & Raney, 1996) do not.

In summary where we fixate during reading is primarily determined by low level factors, such as word length and the location of the prior fixation. While there are some small influences of cognitive factors such as frequency and predictability on whether or not a word will be directly fixated, these are fairly modest when compared to the effects of length and launch site. Fixation positions within a word are influenced exclusively by low-level factors, with no influence of factors such as frequency.

1.2.1. The Perceptual Span and Parafoveal Processing

While a large proportion of the processing of a word takes place in foveal vision, it is not the case that encoding only begins upon direct fixation of the word. Rather, a word is often partially processed on a fixation on a prior word. The parafoveal information extracted from this fixation is then carried over and integrated with foveal information that is available when the word is fixated. Furthermore, a single word is sometimes fixated more than once, in which case the information extracted during these multiple direct fixations must also be integrated. The process of integrating information extracted across multiple fixations is the focus of the following section.

The fact that a word is often processed over multiple fixations is apparent from studies using the moving window paradigm (McConkie & Rayner, 1975; see also Section 1.1.2). As a reminder, in this paradigm, a window of normal text is set by the experimenter around the current point of fixation, outside of which the text is masked. Research using this technique has shown that in English a window spanning 3-4 character spaces to the left and 14-15 characters to the right of fixation is large enough for reading to proceed uninhibited (McConkie & Rayner, 1975; McConkie & Rayner, 1976). Thus, it is clear that readers are able to extract information from 14-15 characters into the parafovea. However, while readers are able to extract information 14-15 characters into the parafovea, the average saccade tends to move the eyes 7-9 characters forward (Rayner, 1998). Therefore, readers usually have overlapping perceptual spans across two fixations, meaning that the same word is often available for processing across multiple fixations. It is clear from this that readers are often integrating information across fixations, since their reading speed decreases when the window is smaller than the size of the perceptual span and a word is thus not available for processing across multiple fixations.

While the moving window paradigm can be used to demonstrate that information is processed across multiple fixations, it does not allow us to infer the nature of this processing, or the type of representation which is integrated across fixations. Several theoretical possibilities exist as to why restricting parafoveal information in moving window studies slows down reading. For example, one early theory proposed that purely visual information obtained from the parafovea is stored between fixations, and that new visual information obtained upon direct fixation is added to this visual representation (McConkie & Rayner, 1976b). Pollatsek, Lesch, Morris, and Rayner (1992) proposed an approach based around the idea that phonological coding serves an important role in silent reading, by helping to create a representation of identified words in short term memory. According to this approach a phonological code is obtained for a word seen in the parafovea, which is used to preserve the memory of that word across fixations. A third possibility is that a parafoveal stimulus activates a set of lexical entries on the basis of several abstract word characteristics (e.g. orthography, phonology, morphology, and semantics), and that this activation is carried across multiple fixations. This lexical activation may then lead to the faster identification of a word once it is directly fixated, thus explaining the slowdown in reading when parafoveal information is denied in the moving window paradigm.

In order to discriminate between the possibilities outlined above it is necessary to manipulate specific characteristics of a single word in the parafovea and examine how this affects fixation times on that word. This has been investigated using the boundary paradigm (Rayner, 1975). As already mentioned above, in the boundary paradigm an invisible boundary is set at the end of a pre-target word, and before the eyes cross this boundary a preview string is presented in place of the target word. I have already mentioned that this technique has been used to demonstrate that fixation times on the target word are reduced when readers receive an identity preview of the target word relative to when they are given an incorrect preview. The existence of this *preview benefit* shows that readers extract and process information about the preview string in the parafovea, and then integrate this information with information obtained on the next fixation. By varying the relationship between the preview and the target, it is possible to discover the types of information that are extracted and integrated across fixations, and to discriminate among different explanations regarding the nature of trans-saccadic integration. For example, if preview benefit effects were purely driven by visual overlap between a preview and target word, it would suggest that the integration process is dependent upon the combination of visual information into a single percept. Were it simply the case that a phonological code is used to aid short term memory of a parafoveal stimulus then preview benefits should only be observed for previews that share phonological information with the target word. If, however, it was to be found that preview benefits are determined by a wide range of abstract information about a parafoveal word, including orthography, phonology, morphology, and semantics, it would suggest that the integration process works on the basis of lexical entries (representing multiple linguistic characteristics) that have been partially activated by the parafoveal stimulus. As will be seen, a substantial amount of research supports this third position, with preview effects being documented for types of information that go beyond low-level visual similarity between the preview and target, or a phonological representation in short term memory.

1.2.2. The Integration of Information during Refixations

Not all saccades move the eyes onto a new word. Rather, a reader will refixate approximately 15% of words (Rayner, 1998). Refixations are more likely for longer words, and the most common pattern is for the initial fixation to be made towards the beginning of a word and the second towards the end (Rayner, Sereno, & Raney, 1996). The fact that refixations often follow this pattern suggests they are made partly due to a need to process characters from the end of a word more centrally within foveal vision. As such, it is worth considering the extent to which information from the end of a long word is integrated with that from the beginning of the same word across fixations. Drieghe, Pollatsek, Juhasz, and Rayner (2010) examined this for monomorphemic words (e.g. *fountain*) and unspaced compounds, which are made of two smaller lexemes (e.g. *bathroom*). A variation of the boundary paradigm which was first implemented by Hyönä, Bertram, and Pollatsek (2004) was used, in which the boundary was placed in the middle of the word instead of between words. In Drieghe et al.'s study, the boundary was between the first and second constituent in the unspaced compounds, or between the corresponding letters of a monomorphemic word that was matched on length, overall frequency, and initial bigram and trigram frequency. In the incorrect preview conditions, the final letters of the target word after the boundary were replaced (e.g., *fountaom*, *bathroan*). There was a large effect of this preview manipulation, with readers having gaze durations of 151ms and 146ms longer in the incorrect preview conditions than in the identity conditions as assessed by the gaze duration on the post boundary portion of the target word for the monomorphemic words and the unspaced compounds, respectively. This demonstrates that when a word is fixated twice, the information from the end of the word has already been processed to a considerable degree during the initial fixation on the word, and this information is integrated with information gained upon refixation.

Although preview benefits from the second portion of the target words were similar for the monomorphemic and unspaced compound words, there were differences in the degree to which the preview affected fixations prior to crossing the boundary. That is, fixations on the first half of a monomorphemic word were lengthened by the incorrect information in the second half of the preview; however, this was not the case for the unspaced compound words. The characteristics of a parafoveal letter string affecting fixation durations on the prior, foveal, text is referred to as a parafoveal-on-foveal effect (for a review see Drieghe, 2011 and Section 1.2.7). These parafoveal-on-foveal effects are generally viewed as evidence that information in the parafovea is being processed in parallel with information in the fovea. Otherwise this parafoveal information would be processed too late to affect the fixation duration on the foveal word. The fact that these preview effects were observed for monomorphemic words but not for unspaced compounds indicates that while the former are processed as single units, the latter are processed, at least to some degree, as two independent sub-units, and this in turn affects the rate at which parafoveal information is integrated.

Häikiö, Bertram, and Hyönä (2010) also investigated the integration of information from the second lexeme of an unspaced compound word (in Finnish, all compound words are unspaced), using the within-word version of the boundary paradigm. They varied whether readers received a correct preview of the second constituent or a preview in which all but the initial two letters were replaced. There was a main effect of the preview manipulation on fixation time on the second constituent, and an interaction between the frequency of the compounds and the preview manipulation during fixations prior to crossing the boundary (i.e. for the high frequency compounds, an incorrect preview resulted in longer fixations on the first constituent before the boundary was crossed whereas this was not the case for the low frequency compounds). Häikiö et al. proposed that the high frequency compounds were identified via a single lexical entry, whereas the low frequency compounds were identified as two separate lexemes; thus the incorrect preview was only processed in parallel with the first constituent in the high frequency compounds.

Hyönä, Bertram, and Pollatsek (2004) demonstrated that the frequency of an unspaced compound's first constituent also affects the integration of information from the second constituent across fixations. They manipulated the frequency of the first constituent of the compound word and all but the two initial letters of the second constituent. These initial letters were either incorrect or correct until the saccade crossed from the first constituent to the second constituent. The preview benefit observed during fixations on the second constituent was larger when the first constituent was a low frequency word than when it was a high frequency word. However, there was no evidence of an effect of the letters of the second constituent being incorrect during fixations on the first constituent—even for the compounds with low frequency first constituents. This suggests that the difference in the preview effect on the second constituent was not caused by the second constituent being processed as part of the whole compound. Rather, Hyönä et al. proposed that these effects were driven by the fact that the low frequency first constituents potentially combined with fewer second constituents to form a compound word than the high frequency first constituents. As such, the set of potential second constituents was more constrained given a low frequency first constituent, and thus was processed more efficiently in the parafovea, as a separate lexeme (See Cui, Yan, Bai, Hyönä, Wang & Liversedge, 2013 for a similar argument in processing Chinese compound words). Taken together, Hyönä et al., Drieghe et al. (2010), and Häikiö et al.'s (2010) findings suggest that a number of factors influence the amount of information that is integrated from the second constituent of an unspaced compound across fixations, and the time course of this processing. How highly constrained the second constituent is influences the amount of information that is integrated across fixations. Furthermore, the frequency of the whole compound influences the time course of when this information first has an observable effect on processing. Further research is required in order to extend current understanding of the factors that determine how the second constituents of unspaced compound words are processed in the parafovea.

White, Bertram, and Hyönä (2008) undertook an experiment in Finnish that investigated whether semantic information from an unspaced compound's second constituent is integrated across fixations. In this study, participants were given a preview of the second constituent of an unspaced compound (*vaniljakastike* 'vanilla sauce') while fixated on the first constituent. There were four possible previews: an identity preview (*vaniljakastike* 'vanilla sauce'), a semantically related preview (*vaniljasinappi* 'vanilla mustard'), a semantically unrelated preview (*vaniljarovasti* 'vanilla priest'), and a pronounceable non-word preview (*vaniljaseoklii*). The identity preview led to shorter fixations on both the second constituent and across the whole compound than any of the other preview types. While the semantically related preview provided no benefit relative to the unrelated and non-word previews in either first fixation duration or gaze duration, on either the second constituent or whole compound, there was a benefit in *regression path durations* within the compound (this includes all fixations within the compound from first fixating the second constituent, until a rightwards saccade was made out of the compound). This fairly late effect of the semantic preview suggests that semantic information was extracted from the second constituent of the compound, but was not integrated immediately upon fixating the target constituent. Rather, integration only occurred during the later phases of compound word processing.

In summary, when a word is fixated multiple times, information is indeed integrated across these fixations. This is true for both monomorphemic and compound words, with substantial preview effects being found within both types of word (Drieghe et al., 2010). The extent to which this information is integrated across fixations within a compound word depends on both the frequency of the whole compound word (Häikiö et al., 2010) and the extent to which the first constituent constrains the second (Hyönä et al., 2004).

1.2.3. The Integration of Information from Word *n*+1

1.2.3.1. Orthographic Codes

A basic question that we can ask in relation to the integration of information across fixations is whether the information that is integrated is based entirely on the visual form of the words or is based on abstract linguistic information that is derived from the orthography of the words. Studies have addressed this by examining the effects of changing the visual characteristics of text across saccades while holding letter information constant. McConkie and Zola (1979) had participants read text in which words were written in alternating case, with the case of each letter changing during saccades (e.g. *ReD* -> *rEd*). This manipulation changed the visual information between fixations, while keeping the letter identities constant. There was no slowdown in reading when the case changed across fixations relative to a condition with no display changes, indicating that the integration of information is not restricted to visual forms. Similarly, Rayner, McConkie, and Zola (1980) showed that participants were no slower at naming a target word when case changed across fixations as opposed to staying the same.

While changing the case of the letters across fixations does not have a significant effect on reading, other work has demonstrated an effect of the visual similarity between a preview and target word. In a meta-analysis of studies in English using the boundary paradigm, Hyönä, Bertram, and Pollatsek (2004) showed that using visually similar replacement letters (e.g. *b* and *d*) results in smaller preview effects relative to an identity condition (15ms in gaze duration, on average) than using visually dissimilar letters (e.g. *p* and *s*; 41ms effect on average). These findings suggest that the orthographic information that is integrated across fixations is in the form of abstract letter identities, which have been activated by low-level visual features. Since visually similar letters will co-activate each other due to shared features (e.g. the vertical ascender in *d* will activate *d*, *b*, and *h*), previews with similar letters will activate the target to a greater extent than previews without similar letters. This also explains why case changes across fixations do not affect reading. While low-level features changed across fixations in these studies, letter identities did not. Thus, a (case-independent) letter representation would have been activated by the features of its lower-case form on one fixation and the features of its upper-case form on the next (and vice versa). As long as letter activation was carried across fixations, it would not have mattered whether this activation was due to the same low-level features on all fixations, or different features on each fixation.

There is evidence that letters from different positions within a word are not equally important when information is integrated across fixations. Inhoff (1989a) gave previews of 6-letter words in which the whole word (e.g. *survey*), the initial trigram (e.g. *surxxx*), the final trigram (e.g. *xxxvey*) or nothing (e.g. *xxxxxx*) was available. Furthermore, the reading direction was varied (e.g., *a recent survey* vs. *survey recent a*), with participants reading from right to left in the latter condition, in order to ensure that any letter position effects were not due to visual acuity. The initial trigram led to slightly, though not significantly, greater facilitation than the final trigram (16ms vs. 12ms in first fixation duration) regardless of reading direction. Furthermore, when visually dissimilar replacement letters were used instead of *x*s Inhoff found that the final trigram alone no longer provided a significant preview benefit, whereas parafoveal availability of the initial trigram still led to a significant benefit of 6ms. Briihl and Inhoff (1995) further investigated this issue by varying the number of correctly previewed letters and their position in a word, and found that previewing external and initial letters was significantly more facilitative than previewing internal letters. One probable reason for the greater benefit of external letters is reduced crowding relative to internal letters, due to being located next to a space. Briihl and Inhoff also found that previews of both final and initial letters together did not facilitate processing significantly more than previews including only initial letters, suggesting that final letters do not play a particularly important role in trans-saccadic integration. However, in both studies, whole word previews were more facilitative than would have been expected had the effect of each extra letter been additive. This suggests that the letters were parafoveally encoded as part of a whole word, and mutually reinforced each other's activation.

While word-initial information in English is given preferential treatment in trans-saccadic integration, this does not generalize to Chinese. Rather than consisting of a string of letters representing a phonological code, Chinese characters are made up of a number of strokes which form sub-units known as radicals. Many characters consist of more than one radical, and the majority of these characters contain a radical that carries phonological information and another that carries semantic information. While these radicals contain this abstract information, the relationship between a character and its radicals is not always strong, with, for example, the pronunciation of only 30% of phonetic radicals corresponding to that of the full character (Zhou & Marslen-Wilson, 1999). As such, two characters with the same phonetic radical may be pronounced differently. Clearly linguistic information is orthographically coded in Chinese in a vastly different way than in English, and so may be integrated differently across fixations. Liu, Inhoff, Ye, and Wu (2002) conducted a boundary study in which the preview and target shared orthographic information via a) the semantic radical, b) the phonetic radical, c) stroke information while sharing neither radical, or d) shared no orthographic information. Liu et al. found that participants gained a significant preview benefit given an overlapping phonetic radical, but not from the other conditions. This effect was observed regardless of whether the target and preview character were phonologically similar. The phonetic radical typically appears on the right side of a Chinese character. Thus orthographic preview benefit is driven by character-final information in Chinese and word-initial information in English. One possible reason for this is that parafoveal orthographic information is used to initiate lexical access, and that the optimal information for this differs across languages. In English the initial letters of a word may be more useful, in part due to their importance in generating a phonological code. However, Liu et al. argued that in Chinese the phonetic radical is more useful for two reasons. First, they claimed that it is the smallest orthographic unit that is always represented in the character lexicon, with it forming a character in isolation. They also claimed that the phonetic radical provides more discriminative information with which to select character candidates from the lexicon. Thus, while different orthographic information is integrated to differing extents in each language, the time course of processing appears to be driven by the underlying principle of what information is most optimally used to initiate lexical access.

As well as investigating how letter identity information is integrated across saccades, researchers have also examined letter position encoding in the parafovea. Johnson, Perea, and Rayner (2007) provided readers with parafoveal previews in which two letters had been transposed (e.g. *loewr* as a preview of *lower*) or substituted (e.g. *loanr*), finding that the transposed letter previews were more facilitative than the substituted letter previews. Johnson (2007) found that this effect endured even when the transposition was made between non-adjacent letters (e.g. *flower* to *flewor*). Johnson and Dunne (2012) presented participants with previews that varied in whether letters were transposed or substituted, and whether they created a non-word or a word which was orthographically similar to the target (e.g., *besat*, and *beats* as transposed letter previews and *berut*, and *beach* as substituted letter previews for the target word *beast*). Preview effects were driven exclusively by the extent of orthographic overlap between the previews and the targets, such that the two transposed letter previews resulted in shorter fixations on the target word than the two substituted letter previews. There was no significant difference between whether the preview was a word or non-word. This study provided further evidence for the transposed letter effect during reading. Furthermore, these findings suggested that processing in the parafovea does not typically proceed to the later stages of lexical processing, during which lexical candidates compete by inhibiting the activation of orthographically similar words. If this had occurred, the word previews should have led to smaller preview benefits than the non-word previews. Together, these studies show that the identity of a letter maintains activation across fixations independent of position. However, this is not to say that letter position *per se* is not important. Clearly it is, since the identity preview always provided reliably more benefit than the transposed letter previews in all of these studies.

The studies I have discussed in this section demonstrate that information about both letter identity and letter position is integrated across fixations. The importance of letter identity information is weighted in relation to a letter's position within a word, and this factor has a differential influence across orthographies.

1.2.3.2. Phonological Codes

One reason for the greater importance of word initial letters in preview benefit may be their role in generating a phonological code to initiate lexical access. Accordingly, it might be expected that an element of such a code might also be taken from the parafovea and integrated with the phonological codes extracted when the word is fixated. In the following section, I consider a series of studies that examined whether phonological codes are integrated across fixations and the nature of these representations.

One way in which phonological processing has been investigated is through the use of homophones in preview studies. Homophones are two words that are spelled differently but pronounced the same. Pollatsek, Lesch, Morris, and Rayner (1992) used the boundary paradigm and presented participants with homophone previews (e.g. *beach* as a preview for *beech*) or orthographic control previews (e.g. *bench*). Participants gained a greater preview benefit from the homophones than the controls. Thus, these results suggest that the overlapping phonological code was integrated across fixations. Chace, Rayner, and Well (2005) replicated this effect, but only in skilled university aged readers, with less skilled university aged readers showing no preview effects. Bélanger, Mayberry, and Rayner (2013) extended the finding by manipulating the relative frequency of the homophone preview and target (i.e., the higher frequency word of the homophone pairs was the preview in half the trials and the target in the other half). Participants gained a phonological preview benefit from the high frequency preview but not from the low frequency preview.

While the above studies demonstrate that readers integrate phonological codes across fixations, it is unclear whether this is driven by addressed or assembled phonology. That is to say, the reader may either gain access to the phonological code via the identification of a complete orthographic representation (a lookup process) or through the use of grapheme-phoneme correspondence rules to assemble a phonological code. Miellet and Sparrow (2004) investigated this in French by giving participants non-word homophone previews (e.g. *maizon* as a preview for *maison*) or orthographic controls (e.g. *mailon*). Despite the homophone preview being a non-word, it facilitated reading. The fact that this effect occurred when the preview strings were non-words suggests that the benefit comes from assembled phonology, since there is no stored lexical representation via which a phonological code might be accessed (for evidence of English readers gaining a phonological preview benefit from non-words see Ashby, Treiman, Kessler, & Rayner, 2006). However, the fact that Bélanger et al. (2013) observed an influence of word frequency on phonological preview effects suggests that readers do sometimes retrieve, as opposed to assemble, a phonological code, with it being possible to extract this information more rapidly from a high frequency word than from a low-frequency word. Thus, depending on the circumstances, readers may make use of either addressed or assembled phonology. Further research is needed to determine the factors affecting which route a reader takes to obtain the phonological code of a parafoveal word.

The studies discussed in this section up to this point all manipulated phonological overlap at a whole word level. Other studies have examined the integration of more fine-grained phonological information within a word. Ashby and Rayner (2004) examined the role of syllabic structure by giving participants previews of words with either a consonant-vowel-consonant (e.g. *concave*) or consonant-vowel (e.g. *device*) initial syllable. A space manipulation was also used so that previews either preserved (e.g. *de_pxw* for device) or violated (e.g. *dev_px*) this structure. Participants remained fixated on the target word for less time when the preview maintained the structure. This was true even for the words with a CV initial syllable, despite the incongruent preview providing more orthographic information. Thus, phonological information at the level of syllables is integrated across fixations, and having these syllables clearly visually delimited in the parafovea may facilitate subsequent processing to a greater extent than a larger number of letters which do not maintain syllabic structure. This suggests that word initial letters may be more facilitative partly because of their role in generating a phonological code. Fitzsimmons and Drieghe (2011) demonstrated that the extraction of a word's syllabic structure in the parafovea must occur rapidly. In this study either a monosyllabic or a disyllabic word matched on word length, frequency, predictability, number of orthographic neighbours and mean bigram frequency was embedded into a sentence. The monosyllabic word was skipped more regularly than the disyllabic word. On the assumption that the parafoveal word's syllabic structure influenced where the next saccade was targeted, this indicates that this information was extracted early enough during parafoveal processing for it to influence saccadic targeting.

Ashby et al. (2006) investigated whether vowel information is integrated across fixations by contrasting vowel concordant and vowel-discordant previews (e.g. *cherg* and *chorg*, respectively, as previews for *chirp*). Vowel concordant previews were more facilitative, even when the vowel's pronunciation needed to be modified by subsequent consonants to be concordant (e.g. *raff* as opposed to *rall* as a preview for *rack*). Thus, this study demonstrated that individual vowel sounds are also integrated across fixations.

The nature of alphabetic languages means that there is a relatively direct link between orthography and phonology in that letters link reasonably reliably to certain phonemes. This is not true for a character based language, such as Chinese. In Chinese, similar-looking characters often have different pronunciations, and homophonic characters may be entirely visually distinct (Hoosain, 1991). Furthermore, as mentioned above, Chinese characters contain a phonetic radical, which in some cases represents the character's phonology, but in other cases contains phonological information which does not match the character's pronunciation. Tsai, Lee, Tzeng, Hung, and Yen (2004) investigated whether Chinese readers integrate phonological information across fixations despite the deeper orthography, and whether the relationship between the phonetic radical and whole character influences this process. Participants were presented with homophonic previews and orthographic control previews. Half of the target characters were pronounced in the same way as other characters sharing the same phonetic radical (i.e. high consistency) and the other half were not (i.e. low consistency). For high consistency targets a phonological preview benefit was observed in both first fixation and gaze duration measures, whereas for low consistency targets the effect was only observed in gaze durations. Clearly, readers of Chinese integrate phonological information across fixations, and this information is extracted from both the whole character and the phonetic radical.

We have seen that phonological information is integrated across saccades both in English and in Chinese where there is a far less clear relationship between orthography and phonology. Furthermore, phonological information is extracted at both the whole word level, the character level, and from sub-units such as syllables and radicals. While we have discussed English as having a fairly direct link between orthography and phonology in comparison to a language such as Chinese, this relationship is less consistent in English than in many other alphabetic languages. As such, future work on parafoveal phonological processing should perhaps focus on these other alphabetic languages with more regular coding schemes more. It may be, for example, that in these languages (e.g., Spanish) even less skilled readers would show evidence of integrating phonological codes across fixations, unlike less skilled readers of English (Chace et al., 2005).

1.2.3.3. Morphological Codes

A further form of information that may be integrated across fixations relates to a word's morphology. Often words can consist of more than one morpheme, and therefore, a word's constituent morphemes may be used to guide lexical access to the whole word form (e.g. *cowboy* may be identified via the lexical entries for *cow* and *boy*). Given this, readers may decompose a parafoveal word into its constituent morphemes, and integrate these units across fixations. If this were the case, then a clearly defined parafoveal morphological unit could impact on subsequent fixations downstream in reading (Lima, 1987).

Several studies have examined this possibility in English (Inhoff, 1989b; Juhasz, White, Liversedge, & Rayner, 2008; Kambe, 2004; Lima, 1987). Researchers have taken the approach of using the boundary technique to provide parafoveal previews to either multimorphemic words (e.g. *revive*, *cowboy*) or monomorphemic control words (e.g. *rescue*, *carpet*) where the previews show a plausible morphemic unit (e.g. *reXXXX*, *carXXX*). The logic behind this manipulation was that a clearly delimited morphological sub-unit might allow participants to initiate lexical access of the word on this basis. For true multimorphemic words this should be facilitative, since the sub-unit would be represented as part of the target word's morphological structure. On the other hand, for the monomorphemic control words, there should be no advantage beyond an orthographic effect. The results of these studies generally suggest that morphology is not extracted in the parafovea, there being no difference between the preview effects for multimorphemic and control words. Both Lima (1987) and Kambe (2004) observed no effect for prefixed words (e.g. *revive*, *dislike*). Lima found no beneficial effects of providing just the prefix (e.g. *disxxxx* for *dislike*) of a multimorphemic word relative to a control word, and Kambe observed no effect of giving either the prefix or the stem (e.g. *xxxlike* for *dislike*). Thus, information about prefixes and affixes does not seem to play a role in trans-saccadic integration during English reading. Inhoff (1989b) found a similar pattern of results for words consisting of two morphemes that can stand alone as words (e.g. *cowboy*). Finally, Juhasz et al. (2008) removed a letter from both compound (e.g. *sawdust*) and monomorphemic (e.g. *lettuce*) words in a position that either preserved (e.g. *saw ust, let uce)* or violated (e.g. *sawd st, lett ce*) a morpheme boundary. The preview that preserved the morpheme boundary did not result in faster processing than the preview that violated this boundary, regardless of the type of word. This suggests that participants did not attempt to process the individual morphemes prior to direct fixation.

These studies provide little evidence that English words are decomposed into their constituent morphemes in the parafovea. Similarly, effects have not been observed in Finnish, a language in which spatially concatenated compounds are very common. Bertram and Hyönä (2007) gave participants previews of Finnish compounds that had a short (3-4 letters) or long (8-11 letters) first constituent, and were on average 12 letters long. The preview consisted of the whole compound or just the first three or four letters. This comprised all of the short first constituents, but not of the long first constituents. Were morphological sub-units being integrated across fixations, then a smaller difference between the two preview conditions for the compounds with short first constituents should have occurred, since participants should have gained a greater morphological benefit from the partial preview for the words with short first constituents. However, no interaction was observed between the preview type and first constituent length, suggesting that parafoveally available morphological information was not being used to initiate lexical access.

While morphological units may not be integrated across fixations in English and Finnish, morphological preview effects have been found in Hebrew (Deutsch, Frost, Pelleg, Pollatsek, & Rayner, 2003; Deutsch, Frost, Pollatsek, & Rayner, 2000; Deutsch, Frost, Pollatsek, & Rayner, 2005). In Hebrew, all verbs and most nouns and adjectives consist of two morphemes. One morpheme is the root, which represents the semantic nature of the word and consists of a series of three consonants. The other is the word pattern, which modifies the root by giving the word its class (i.e. noun, verb, adjective, etc.) and other characteristics. Of these two morphemes the root is more important to word meaning, and thus in Hebrew words there are three letters which provide more useful information than the rest of the letters. The two morphemes are interwoven, rather than concatenated. For example, the root morpheme חבר and the word pattern ת___מ take the form of a word with interwoven constituents like תחברמ rather than a concatenated format like תמחבר. Word patterns' structures are highly constrained, such that they can only begin with certain consonants, and each letter imposes a set of transitional probabilities on subsequent letters. Consequently, it is possible for readers of Hebrew to rapidly determine which letters belong to the word pattern, and which belong to the root morpheme. In sum, within Hebrew words there are several letters which carry more useful semantic information than the others, and these letters are more easily located within and thus extracted from a word. Due to this, readers of Hebrew may be able to rapidly decompose a word into its constituent morphemes in the parafovea, and then integrate these morphemes across fixations.

Deutsch et al. (2000) first investigated whether the root morpheme is integrated across fixations in Hebrew using a naming paradigm. In this study, an isolated preview of a target word was presented in the parafovea. This preview was either the target word (e.g. תחברמ), the three letters of the root morpheme (e.g. חבר), an orthographic control (e.g. מבת), or an X-string. Participants gained a benefit from the morphological preview relative to the orthographic control, such that they named the target more quickly upon fixating it. Deutsch et al. (2003) extended this finding by showing that a morphological preview benefit is obtained during sentence reading using the boundary paradigm, and when the letters of the root morpheme had to be extracted from the letters of the word-pattern, rather than being presented as an isolated unit. One preview was morphologically related to the target, in that it included the target word's root morpheme within an alternative word pattern. This provided a preview benefit relative to an orthographic control, which shared the same number of letters with the target but was derived from a different root. Participants had clearly extracted the root morpheme in the parafovea, and used this to guide lexical access.

Deutsch et al. (2005) also investigated whether the morphological code of the word pattern is integrated across fixations. This was examined for both verbal patterns (i.e. word patterns that combine with the root to form a verb) and nominal patterns (i.e. word patterns that combine with the root to form a noun). An important difference between these two types of word patterns is that while the verbal patterns possess properties that may guide lexical access, the nominal patterns do not. Specifically, nominal patterns do not have precise semantic characteristics, and the frequency of most nominal patterns is low in comparison to the frequency of the verbal patterns. Deutsch et al. showed that it is possible to gain a morphological preview benefit from a preview consisting of the word-pattern in an alternative word in the case of verbs, but not nouns.

In summary, Hebrew readers decompose words into their constituent morphemes in the parafovea, and then integrate this information (usually) on the following fixation on the word in order to aid lexical identification. There is clearly a difference between parafoveal morphological processing for readers of Hebrew and readers of English and Finnish. The cross-linguistic difference that may most plausibly account for this is the speed with which it is possible to extract individual morphemes in the parafovea. In Hebrew there are strict rules governing which letters within a word can belong to each morpheme. This is not the case in English, with there being relatively few constraints upon where one morpheme ends and another begins. Indeed, the existence of the monomorphemic control words used in the English studies demonstrates this, with it being possible for *re* to either be a prefix or two letters in a monomorphemic word. Thus, readers of Hebrew have stronger cues with which to reliably morphologically decompose words than readers of English, and these cues may partially account for differences in the parafoveal extraction of morphological units.

1.2.3.4. Semantic Information

Over the past several decades the predominant view has been that semantic information is not integrated across fixations, due to early findings from studies conducted primarily in English. Rayner, Balota, and Pollatsek (1986) presented participants with previews of a target word (e.g. *father*) that were semantically related (e.g. *mother*), orthographically similar (e.g. *fatlon*) or unrelated (e.g. *circle*). The semantically related previews provided no benefit, suggesting that semantic information was not carried over to subsequent fixations (see Rayner, Schotter, & Drieghe, 2014 for a replication). A similar pattern of results was found in a gaze-contingent naming study (Rayner et al., 1980). Further evidence against semantic information being integrated across fixations was found by Altarriba, Kambe, Pollatsek, and Rayner (2001). In this study, Spanish-English bilinguals read sentences with previews that were translations of a target word which were either orthographically similar (e.g. *crema* as a preview for *cream*) or dissimilar to the target (e.g. *fuerte* as a preview of *strong*), orthographically similar words in the opposite language that were not translations (e.g. *grasa* as a preview for *grass*), or an unrelated word in the opposite language (e.g. *torre* as a preview for *cream*). Since the translation shared a semantic representation with the target word, it was hypothesized that significantly more preview benefit might occur for the translation preview than the orthographically similar non-translation if semantic information was integrated across fixations. However, the amount of preview benefit was primarily driven by orthography, and not semantics. This study offers little support for the view that semantic information is integrated across fixations.

Research conducted on semantic preview benefit in Finnish also suggests that this information is not integrated across fixations. Hyönä and Häikiö (2005) gave participants parafoveal previews that were either correct (e.g. *pentu* ‘cub’), emotionally arousing (e.g. *penis*), or neutral (e.g. *penni*, ‘penny’). They hypothesized that if readers extracted semantic information from these previews then there would be disruption to reading in the emotional condition, due to the possibility that this information would be arousing enough to disrupt processing. However, there was no effect of the emotive content of the preview.

Although these studies suggest that semantic information is not integrated across fixations, recent evidence suggests that this is not necessarily the case. Reliable semantic preview effects have now been observed in several studies of Chinese reading. In Chinese the majority of characters include a semantic radical, and therefore, there is a more direct link between the orthography and semantics of a word than in alphabetic languages. This makes it more likely that semantic information can be extracted in the parafovea, and then integrated on the next fixation. Yan, Zhou, Shu, and Kliegl (2012) examined whether semantic information from both the radical and character level is integrated across fixations. Participants were given an unrelated preview, and two different types of semantically related previews. One of the semantically related previews was semantically transparent, in that the meaning of the character was congruent with the meaning of the semantic radical, whereas the other was opaque. None of the previews contained the same semantic radical as the target character, and so any preview benefit could not be due to orthographic confounds. Yan et al. found that both types of semantic preview led to shorter reading times than an unrelated preview, with the semantically transparent preview leading to a larger benefit in gaze duration than the semantically opaque preview. This pattern of results demonstrates that semantic information from both the whole character and the radical is activated in the parafovea, and that both types of semantic information are then integrated with semantic information extracted from the target character upon fixation. This can be seen from the fact that semantic overlap between the preview and target character reduced target fixation durations, and that there was a greater effect when the preview's semantic radical and the target character also shared semantic information. Furthermore, Yan et al. observed larger semantic preview effects when fixation times on the pre-boundary word were longer (see Hohenstein & Kliegl, 2014 for a discussion of this effect).

Semantic preview effects have also been observed in German. Hohenstein, Laubrock, and Kliegl (2010) found effects in German using parafoveal fast priming. In this technique, a non-word preview of the target word is present until readers make a saccade over an invisible boundary prior to the pre-target word. As a saccade is made onto the pre-target word, a display change is triggered. In the Hohenstein, et al. (2010) experiment, this led the target word to change to either a semantically related or an orthographically matched preview for a set amount of time before becoming the target. The amount of time the parafoveal preview was available for was varied. At short prime durations (e.g. 35, 60, and 80ms) there was no semantic preview benefit. At a longer prime duration (125ms) there was a significant semantic preview benefit of 24ms. Furthermore, there was a change in this pattern of effects when the target word was made more salient via being presented in bold. Here a significant semantic preview benefit of 18ms was found at the 80ms prime duration, but no facilitation was found for the 125ms prime. The authors claimed that this was due to semantic information being facilitative only up to a certain moment, beyond which the orthographic mismatch overrides the effect. Some caution may be necessary in interpreting these results as it is not entirely clear how the visual changes that occur in the fast priming technique influence attentional allocation during reading.

Hohenstein and Kliegl (2014) found further evidence for semantic preview benefit in German using the standard boundary paradigm. They found that a semantically related preview (e.g. *Schädel* ‘skulls’ as a preview for *Knochen* ‘bones’) was more facilitative than an unrelated preview that shared the same amount of orthographic information with the target word (e.g. *Stiefel* ‘boots’). This effect was reliable across fixation time measures over three experiments and averaged 26ms in gaze duration. Furthermore, the effect endured regardless of whether the target noun was capitalised or not (in German, nouns are capitalised). This is important since it may be easier to extract parafoveal semantic information for nouns in German since the capitalization may give readers a salient cue to the syntactic class of the parafoveal word, allowing for more processing resources to be allocated to that word than might otherwise be the case. Furthermore, there was an effect of pre-target fixation duration that was similar to that reported by Yan et al. (2012), such that there was a greater semantic preview benefit following longer fixations on the pre-target word.

The final study we will consider in this section is that of Schotter (2013). In this study investigating reading of American English, participants were given two different types of semantically related previews. The first type (e.g. *rollers* as a preview for *curlers*) was highly related to the target (7.5 on a 9 point rating scale in a norming study) and maintained the sentence meaning (7.2 on a 9 point rating scale). The second type (e.g. *styling*) was less semantically related (5.6) and maintained the sentence meaning to a lesser extent (4.9). Unrelated previews (e.g. *suffice*; 2.4 and 1.9 on the rating scales) were also included. All three previews shared a similar amount of orthographic information with the target. Relative to unrelated previews the highly related previews led to shorter fixation durations on the target word (16 and 19ms in gaze durations across two experiments). There was no benefit from less semantically related previews. Furthermore, the extent to which the preview changed the meaning of the sentence predicted fixation times on the target word. Schotter argued that this suggests the lack of effects in English in prior studies arose because the semantic relationship between the preview and the target word did not preserve meaning to the same degree that her stimuli did. For example, Rayner et al. (1986) used target-preview pairs such as *father-mother*, *ocean-river*, and *sick-well*, which while semantically related to each other, did not necessarily share the same meaning.

In sum, the evidence regarding whether semantic information is integrated across fixations is currently mixed. Some studies have failed to show clear effects, while other studies do appear to show effects often under specific experimental circumstances. It is not possible at present to provide a coherent explanation of the current state of this aspect of processing — in some senses it is quite contradictory. Further research is necessary in order to gain a clearer understanding.

1.2.4. The Integration of Information from Word *n*+2

The preceding sections have all focused on how various types of information about the upcoming word (*n*+1) are integrated across fixations. Recently, however, research has begun to investigate whether information from word *n*+2 is also integrated across fixations (Angele & Rayner, 2011; Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Kliegl, Risse, & Laubrock, 2007; Rayner, Juhasz, & Brown, 2007; Risse & Kliegl, 2012).

To investigate the integration of information from word *n*+2 across fixations, researchers have manipulated the preview of a word while it is two words to the right of fixation, with the preview changing to the target as a saccade is made onto the pre-target word (word *n*+1). Any effect of this manipulation would suggest that readers are extracting information from word *n*+2 when it is in the parafovea, and integrating this information during subsequent fixations. Rayner et al. (2007) presented participants with either a correct or incorrect preview of a target word, and manipulated whether the boundary was directly before the target word, or directly before the pre-target word. As such, the incorrect preview was either visible as word *n*+1 or word *n*+2. The preview manipulation only had an effect when the preview was visible as word *n*+1. Thus, Rayner et al. did not observe evidence for the integration of information from word *n*+2. Kliegl et al. (2007) further investigated this issue. In their study, word *n*+1 was always three letters long, thus ensuring that the preview of word *n*+2 was as close to central vision as was reasonably possible. Furthermore, they tested for effects of the *n*+2 preview on fixation times on both word *n*+1 and *n*+2. While the *n*+2 preview did not affect fixations on word *n*+2, it did affect fixations on word *n*+1, suggesting that information from word *n*+2 was extracted (see Risse & Kliegl, 2012, for a discussion and test of why this effect appeared on word *n*+1). Angele et al. (2008) orthogonally manipulated previews of word *n*+1 and *n*+2 and ensured that word *n*+1 was always at least four characters long. They found that while there were reliable *n*+1 preview effects, there were no effects of the *n*+2 preview. The posited reason for the discrepancies across studies is the length and processing difficulty of word *n*+1. When word *n*+1 exceeds three characters it is more difficult to process, and therefore word *n*+2 is less likely to be processed before a saccade is made across the boundary. Furthermore, even when word *n*+2 is processed, information extraction occurs less efficiently, since it is further into the parafovea.

Angele and Rayner (2011) manipulated whether readers received identity or non-word previews of a three letter word *n*+1, and a word *n*+2 which was on average seven letters long. While *n*+2 preview effects were found when there was an identity preview of word *n*+1, there was no effect when it was a non-word. Thus, when word *n*+1 cannot be lexically processed (due to it being a non-word) information from word *n*+2 does not appear to be integrated.

More recently Cutter, Drieghe, and Liversedge (2014; see also Chapter 2) found an *n*+2 preview effect even when word *n*+1 was long. In this study, word *n*+1 (e.g., *teddy*) was on average 5.65 letters long, and formed a spaced compound (e.g. *teddy bear*) with word *n*+2 (e.g. *bear*). Participants were given either a correct preview of both constituents, of only the first constituent, of only the second constituent, or of neither constituent. When the first constituent was correct, participants gained a sizeable *n*+2 preview benefit, such that gaze durations on word *n*+1 were 27ms shorter when there was a useful preview of both constituents, rather than just the first. This demonstrates that while *n*+2 preview effects are not typically observed given a long word *n*+1, this can be modulated by the extent to which word *n*+2 forms a single multi-word unit with word *n*+1. Furthermore, it shows that the absence of *n*+2 preview effects in prior studies was not due to visual limitations.

In summary, there is evidence for information from word *n*+2 being extracted, and arguably, integrated across fixations in English, but only under specific circumstances. The studies reviewed suggest that word *n*+1 must be short and easy to process for information from word *n*+2 to be extracted and integrated across fixations. Furthermore, even when such effects are observed they are small (e.g. 7 to 20ms) when compared to effects of word *n*+1 (e.g. 20 to 50ms). The one exception to this is when word *n*+2 was part of a spaced compound, an issue that we will return to below.

1.2.5. Modulation of Parafoveal Processing

So far, I have discussed the extent to which information is integrated across fixations as if this is an invariant process. However, several factors have been shown to modulate this process. The first is foveal load, and the second is the extent to which the foveal and parafoveal word can be considered a single unit.

Foveal load refers to the difficulty of processing on any particular fixation. When the currently fixated word is difficult to process, then foveal load is high. It has been argued that increased foveal processing load results in reduced parafoveal processing (Henderson & Ferreira, 1990; White, Rayner, and Liversedge, 2005), thus reducing the extent to which information may be integrated across saccades. Henderson and Ferreira manipulated foveal load via either a word frequency or a syntactic manipulation and presented participants with a correct or incorrect preview using the boundary paradigm. Significant effects of the preview type were only observed when the foveal load was low. The effect of the foveal word's frequency on preview benefit has also been observed by White et al. (2005).

While several studies have shown that foveal load modulates the parafoveal preview benefit, research by Drieghe, Rayner, and Pollatsek (2005) suggests that this is not always the case. In this study, foveal load was varied using the same frequency manipulation as in earlier studies, and participants were given a preview of a three-letter target word. However, the size of the preview benefit was the same for the high and low foveal load conditions. Drieghe et al. proposed that the absence of an interaction may have been due to the short parafoveal words being processed differently from the longer parafoveal words used in other investigations of foveal load. However, it is unclear why the length of a parafoveal word would determine the extent to which foveal load influences parafoveal processing. More work is required to further explore this effect.

A second factor that influences how far into the parafovea information is extracted from and then integrated across fixations is the degree to which the foveal and parafoveal text is unified spatially and linguistically. This is an issue that has been touched upon throughout the current chapter. In terms of spatial unification, a larger preview benefit is observed when the preview is of the end of the fixated word (e.g. 151ms in gaze durations on the second half of a word in Drieghe et al., 2010) than when the preview is of the word to the right of fixation (e.g. an average of 41ms for dissimilar letters in Hyönä et al., 2004). Even less of an effect is observed from previews of word *n*+2, with the literature only finding effects of between 7 and 20ms. The one exception to this was Cutter et al.'s (2014) study, in which a 27ms effect was observed in gaze duration on word *n*+1 when word *n*+2 formed a spaced compound with word *n*+1. This effect suggests that whether two physically separated parafoveal words form a single lexical unit or not influences the amount of information integrated across fixations.

The results of several studies suggest that the lexical unification of information within a fixated word also influences the extent and time course of the integration of information from the end of this word. As discussed in the section on within-word integration, it has been found that people differentially integrate information from the end of unspaced compounds and monomorphemic words (Drieghe et al., 2010). Furthermore, Häikiö et al.'s (2010) study suggested that information from the end of unspaced compounds was integrated differently depending upon whether the compound was identified as a single lexical unit or two separate lexemes. Häikiö et al. showed that when an unspaced compound was identified as a single lexical unit, incorrect information at the end of the second constituent was integrated early enough to affect fixations on the first constituent. This was not the case when the unspaced compounds were processed as two separate lexemes. Thus, this research suggests that the time course in which information is integrated across fixations is modulated by whether the information in the fovea and parafovea are processed as part of the same lexical unit.

There is also evidence that a greater amount of information is integrated from word *n*+1 when it forms part of a larger unit with the fixated word (Inhoff, Starr, & Shindler, 2000; Juhasz, Pollatsek, Hyönä, Drieghe, & Rayner, 2009). Inhoff et al. examined preview benefit for the second word of spaced compounds (e.g. traffic light, fairy tale, video tape). This study found a considerably larger preview benefit than is usual between words, such that there was a 91ms effect of a dissimilar preview in comparison to the average of 41ms (Hyönä et al., 2004). Furthermore, the manipulation affected fixation times on the first constituent, in a similar manner to preview manipulations within monomorphemic words (Drieghe et al., 2010) and frequent unspaced compounds (Häikiö et al., 2010). Juhasz et al. also found a larger than usual preview benefit for the second constituents of spaced compounds (34ms vs. an average of -7ms for studies using an equivalent level of disruption), although this study did fail to find significant differences between spaced compounds and adjective-noun pairs, for which there was a 21ms effect. The findings of both Inhoff et al. and Juhasz et al. suggest that a greater amount of information may be extracted from a parafoveal word if it forms part of a larger unit with the foveal word. Furthermore, Inhoff et al.'s finding indicated that this parafoveal information may have been integrated earlier than is typical for a parafoveal word that does not form a single unit with the fixated word.

To summarise, several factors have been found to influence the extent to which parafoveal information is integrated across fixations. One is foveal load, with preview benefit effects being reduced when the fixated word is difficult to process. The second factor is the extent to which the information in the parafovea forms a single unit with the fixated word.

1.2.6. A Summary of Parafoveal Preview Effects

We have seen that a large variety of information is integrated across fixations, from both the end of a single word and from a parafoveal word. The integration of information from word *n*+1 operates on the basis of abstract codes for word characteristics such as orthography, phonology, semantics, and, in the case of Hebrew, morphology. There are several interesting cross-linguistic differences that influence the information which is preferentially integrated across fixations. For example, readers of English preferentially integrate word-initial letters, Chinese readers integrate the final radical of the parafoveal character, and Hebrew readers integrate morphological codes. The underlying reason for these differences may well be the extent to which the information allows the reader to initiate lexical access. For readers of English, the phonological code granted by the word initial letters may be most useful, while in Hebrew the root morpheme may provide more useful information for activating appropriate lexical candidates. Finally, in Chinese the final radical may provide more discriminative information to activate a limited set of character candidates. As such, research suggests that information is integrated across fixations on the basis of the most useful information for identifying words in a particular language. While the research shows that a large amount of information about word *n*+1 is integrated across fixations, the same is not true for word *n*+2, with preview manipulations to this word having small effects that only occur under optimal conditions. Finally, the way in which readers integrate information across fixations is influenced both by foveal load and whether the parafoveal text forms a larger unit with either the foveal text or more distal parafoveal text.

1.2.7. Parafoveal-on-foveal Effects

As has been touched upon throughout the above section, information processed in the parafovea may not only affect direct fixations made on that word. Rather, it may also affect fixations made on the prior word while still in the parafovea, known as a parafoveal-on-foveal effect. This is an area of some controversy, with such effects being viewed as potentially discriminating between two different theoretical approaches to eye movement control and lexical processing during reading. The first of these approaches assumes that words are processed one at a time in serial order, with the processing of a word not beginning until all preceding words have been fully identified. The second approach assumes that all words within the perceptual span are processed at once, in parallel. Parafoveal-on-foveal effects are typically viewed as evidence for the latter approach. The rationale for this claim is that if words are processed in serial order then the characteristics of an upcoming word should not be able to affect the processing of the fixated word, since the fixated word should have been fully identified before the processing of the parafoveal word begins. On the other hand, if words are processed in parallel then the characteristics of and information from an upcoming word will be processed at the same time as the fixated word, thus allowing the processing of this parafoveal word to affect the processing of the fixated word. However, such effects may not always be due to parallel lexical processing; rather, feasible explanations exist which are consistent with a serial approach to lexical processing. Before discussing these alternative explanations, I will first briefly summarise research demonstrating parafoveal-on-foveal effects.

Parafoveal-on-foveal effects can be roughly divided into two categories. The first category consists of effects due to the orthography of the parafoveal word. One way in which the orthography of the parafoveal word has been found to affect fixation times on the foveal word is through unusual orthographic information in the parafovea disrupting the processing of the fixated word. Several studies have shown reliable effects of orthographically illegal or unusual information at the start of a parafoveal word in viewing times on the fixated word (Drieghe, Rayner & Polltasek, 2008; Inhoff, Starr, & Shindler, 2000). In these studies the boundary paradigm was used to present either an orthographically illegal (e.g. *pvxforming*) or correct word (e.g. *performing*) in the parafovea. It was found that readers fixated word *n* for longer when it was followed by an orthographically illegal word, suggesting a disruptive influence of the parafoveal information. Furthermore, White (2008) found increased fixation durations on a word when the following word was orthographically unfamiliar (e.g. *crypt*) rather than familiar (e.g. *adder*). It has also been found that orthographic information from a parafoveal word can facilitate the processing of a fixated word. Angele, Tran, and Rayner (2013) found that participants fixated a target word (e.g. *news*) for less time when they manipulated the parafoveal preview of the following word to be identical or orthographically similar (e.g. *niws*) rather than orthographically dissimilar (e.g. *tale, tule, once*). Similar findings were made by Dare and Shilcock (2013) and Inhoff, Radach, Starr, and Greenberg (2000).

The second category of parafoveal-on-foveal effects is lexical. A lexical parafoveal-on-foveal effect would involve characteristics such as the frequency, meaning, or predictability of a parafoveal word influencing fixation durations on the fixated word. Many studies have failed to find any such effects. For example, White (2008) found that while the sub-lexical factor of orthographic familiarity caused a small parafoveal-on-foveal effect, frequency did not. In this study, participants would fixate the word preceding an orthographically unusual and infrequent word (e.g. *quay*) for longer than when the word preceded an equally infrequent but orthographically typical word (e.g. *cove*). However, there were no differences between fixation times on the target word when it preceded a high frequency orthographically typical word (e.g. *town*) rather than the low frequency orthographically typical word. Thus, all differences in this study were driven by orthographic regularity, rather than frequency. Henderson and Ferreira (1993) also failed to find frequency driven parafoveal-on-foveal effects.

Some studies have actually observed lexical parafoveal-on-foveal effects. For example, it has been found that the plausibility of word *n*+1 in the sentential context is able to affect fixation durations on word *n* (Rayner, Warren, Juhasz, & Liversedge, 2004). In this study it was found that if word *n*+1 was completely anomalous within the sentential context (e.g. *carrots* in *he used the pump to inflate the large carrots*) it led to increased fixation durations on word *n*, although only when this fixation was on the last three characters of word *n*. One study that did find evidence for semantic parafoveal-on-foveal effects was conducted by Inhoff, Radach, Starr, and Greenberg (2000). In this study it was found that when word *n*+1 was either identical to or semantically associated with word *n* (e.g. *mother's mother*, *mother's father* with *mother's* as word *n* and *mother*/*father* as word *n*+1) it led to shorter fixation durations on word *n* than when word *n*+1 was not related to word *n* (e.g. *mother's garden*). Furthermore, this effect was not constrained to fixations on the last three letters of word *n*. This suggests that the processing of word *n* was affected by semantic information extracted from word *n*+1, thus constituting a parafoveal-on-foveal effect. However, it should be noted that Angele et al. (2013) failed to observe a similar effect, finding that fixation times on a foveal word (e.g. *news*) were unaffected by a semantically related preview of the parafoveal word (e.g. *tale*) relative to an unrelated or non-word (e.g. *once*, *tule*).

Whilst lexical parafoveal-on-foveal effects have not always been particularly forthcoming in tightly controlled experiments, evidence has been observed in large scale corpus analyses. Kennedy and Pynte (2005) found, using a corpus of 50,000 words, that in cases when the foveal word was short (5 to 6 letters), and thus word *n*+1 tended to be in an area of comparably high visual acuity, a low frequency word *n*+1 led to longer fixation durations on word *n*. This is quite clearly a lexical parafoveal-on-foveal effect. Kliegl, Nuthmann, and Engbert (2006) found a similar parafoveal-on-foveal effect of frequency in a further corpus study in German. Furthermore, they also found an effect of the predictability of word *n*+1, such that a highly predictable word *n*+1 was related to longer fixations upon word *n*. However, as mentioned in Section 1.1.3 there are certain drawbacks of corpus-based analyses, and it may be the case that studies reporting lexical parafoveal-on-foveal effects suffer from these drawbacks. Specifically, due to the uncontrolled nature of these studies it can sometimes be the case that the independent variable under investigation is confounded with a second variable in the corpus, and that it is actually this second variable causing the effect as opposed to the variable that the effect is attributed to.

A recent study by Angele, Schotter, Slattery, Tenenbaum, Bicknell, and Rayner (2015) elegantly illustrated both the discrepancy between corpus-based and experimental approaches to the investigation of parafoveal-on-foveal effects, and the fact that confounding variables can lead to the observation of effects which are misattributed to the lexical processing of a parafoveal word. In this study participants read a large corpus of sentences. During reading, a gaze-contingent moving mask was applied to the text, such that word *n*+1 or *n*+2 would or would not be masked. Analyses were carried out on the majority of words in the text in an analysis consistent with the corpus approach used by Kliegl et al. (2006) and Kennedy and Pynte (2005). Furthermore, they also presented certain sentences in which a specific target word had been manipulated for frequency, and they analysed the effect of this manipulation on the preceding word, in something more akin to a tightly controlled experimental design. In line with the prior research they did not observe any parafoveal-on-foveal effects in their analysis of a single target word, but did observe effects in their corpus analysis. However, these effects were observed regardless of whether or not a parafoveal preview of the upcoming word was available. In other words, the corpus analysis suggested that the characteristics of word *n*+1 influenced the processing of word *n* even when it was not possible for readers to extract these characteristics. This suggests that effects which are attributed to parafoveal processing may actually be due to other factors which are not accounted for in the statistical models used in these corpus studies. Angele et al. suggested that readers may use subtle cues in the text to form expectations about the difficulty of upcoming words, and modulate their reading speed accordingly; assuming that these expectations are correct, then this will lead to a correlation between the difficulty of the upcoming word and fixation times. Whether this particular explanation is correct or not, this study demonstrates that results obtained via a corpus analysis should be treated and interpreted with caution.

The issue of whether parafoveal-on-foveal effects are actually due to the parallel processing of two words is further complicated by the issue of both systematic and random error in saccadic targeting sometimes resulting in a fixation being located on a word other than the one that was being targeted (McConkie, Kerr, Reddix, & Zola, 1988; see Section 1.1.4). Such error in the oculomotor system gives rise to an explanation for parafoveal-on-foveal effects that is consistent with a serial view of processing, known as the mislocated fixations account (Drieghe, Rayner, & Pollatsek, 2008). Often, parafoveal-on-foveal effects have only been found in studies using controlled experimental designs when analyses were restricted to fixations on the final three letters of the foveal word (e.g. Rayner, Warren, Juhasz, & Liversedge, 2004). This leaves open the possibility that these effects are actually due to a saccade intended for the parafoveal word landing on the end of the foveal word. In such cases, attention would still be directed to the intended target, but the fixation would be on the foveal word. As such, the duration of this fixation would be determined by the qualities of the parafoveal word, leading to the observation of parafoveal-on-foveal effects caused by lexical processing of the parafoveal word. Drieghe et al. derived several markers of such a phenomenon and conducted an experiment to test whether parafoveal-on-foveal effects are related to these markers. These markers are that any effects will be independent of the characteristics of the foveal word since this is not the word that is being processed; that such effects will only be seen when fixations are very close to the parafoveal word since the saccade was intended for this word; and that the fixation durations on the foveal word should be correlated with the length of the prior saccade due to the range error causing the mislocated fixation. It was found that the circumstances under which parafoveal-on-foveal effects occurred supported a mislocated fixations account of such effects.

A further issue regarding whether parafoveal-on-foveal effects are in fact due to the parallel processing of multiple words relates to measurement error of where the eye is actually fixated (Reichle & Drieghe, 2014). While eye tracking technology is highly accurate, and can typically give a measurement of which character within a word is currently fixated, there are several factors which can systematically bias this measurement. One such factor is the quality of the calibration. At the start of and during eye tracking experiments participants will typically be required to fixate a series of targets, with the same series of targets being presented twice. This calibration procedure allows the eye-tracker to make an estimate of the correspondence between the recorded image of the eye and where a fixation is located on a display. This estimate will then be used during the study to determine which character within a word a reader is fixating at any one time. If a researcher is competent the calibration should be highly accurate, and thus the fixation times and locations will accurately represent the processing of the word that is supposed to be fixated. If, however, the calibration is inaccurate, the fixation time and location may not accurately represent the processing of the fixated word. Take, for example, a situation in which the calibration will lead to an estimate which suggests that the eye is fixated two characters further to the left than it actually is. In such cases fixations that are located on the first character of one word will be detected as being on the final character of the previous word. Consequently, there will be a proportion of fixations allocated to one word that actually reflect the processing of the upcoming parafoveal word. This measurement error will, at times, lead to the false observation of parafoveal-on-foveal effects. Reichle and Drieghe demonstrated that within the framework of a model of eye movement control that assumes serial lexical processing (see Section 1.3.2) measurement errors can indeed lead to the observation of apparent parafoveal-on-foveal effects. As such, it is worth bearing in mind that these effects may at least be partially driven by measurement issues, rather than the parallel processing of multiple words.

In summary parafoveal-on-foveal effects are highly contentious, both in terms of their theoretical implications and the reliability of their observation. While effects due to orthography are found reasonably reliably, lexical effects are far more inconsistent, with the majority of observations of such effects being observed in corpus analyses, as opposed to tightly controlled experimental settings. In addition to the lack of consistency of some of these effects there are alternative explanations for them that do not depend upon the parallel processing of multiple words.

1.3.1. Computational Models of Eye-Movement Control during Reading

The preceding sections have made it clear that a great deal of knowledge about eye movement control during reading has accumulated over the past several decades. In the next section I will explain how these findings have been implemented in models of eye movement control in order to allow such models to account for when the eyes move, the time course of parafoveal processing, and the location of initial fixations during reading. There are some issues on which multiple models agree, such as saccadic targeting during reading. On the other hand, there is a great deal of debate in other areas, such as the time course of the processing of parafoveal information. Throughout this thesis the assumptions of these models will be taken into consideration and assessed in order to discriminate between the two dominant models, as well as to determine if there are parameters common to both models that require refinement.

While there are many models of eye movement control, there are two contrasting models that currently dominate the debate. These are the E-Z Reader model (Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003; Pollatsek, Reichle, & Rayner, 2006; Reichle, Warren, & McConnel, 2009), and the SWIFT model (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005; Schad & Engbert, 2012). The main point of contention between these two models is the extent to which we process words beyond word *n*, and the time course of this processing. According to the E-Z Reader model, lexical processing is serial, in that word *n* is fully processed before attention moves onto word *n*+1, and the processing of word *n*+1 must be fully completed for processing to begin on word *n*+2, and so on. In contrast to this SWIFT proposes that lexical processing takes place in parallel, such that all of the words that fall in the perceptual span on any given fixation are all lexically processed at the same time. These assumptions will be addressed in more detail below.

1.3.2. The E-Z Reader Model

At a basic level, the E-Z Reader model assumes that attention is allocated to one word at a time, and that the initiation of a saccade is tightly linked to the lexical processing of this fixated word. In the E-Z-Reader model there are two stages of lexical processing. The first is L_1, known as the familiarity check, which involves checking to see if a word's orthographic form seems familiar. The amount of time this stage takes is modulated by the frequency and in-sentence predictability of the word, with the duration of L_1 decreasing as the frequency and predictability of a word increases. Once this stage is completed a saccade program will be initiated, in order to prepare the eyes' movement to the next word. The modulation of the duration of L_1 by a word's frequency and predictability explains the effect of these variables on fixation durations (see Section 1.1.3), with the preparation of a saccade away from an easy word beginning earlier than from a difficult word. The second stage of lexical processing is L_2, which is called lexical access. L_2 is a fixed proportion of L_1, and thus also modulated by a word's predictability and frequency. It is during L_2 that a word's meaning becomes activated, and after L_2 is completed attention is deployed to the next word in the parafovea. At the same time that L_2 is happening, a saccade to the parafoveal word is programmed. It is important to note that the time to program a saccade is a constant in the model, while the duration of L_2 is modulated by the frequency and predictability of the attended word. Consequently, a saccade is not always ready to execute by the time the fixated word has been fully identified. Due to this word *n*+1 will be processed while still in the parafovea until a saccade is ready to be made to this word. It is this decoupling between the deployment of attention and the execution of a saccade that explains the various types of preview benefits for information from word *n*+1 discussed in Section 1.2.3. This also explains the effect of foveal load on the size of these preview effects; when a word is lower frequency L_2 takes longer to complete, and thus attention does not spend as long on word *n*+1 prior to a saccade being launched. I will defer discussion of how this model explains the *n*+2 preview effects discussed in Section 1.2.4 until Chapters 2 and 3, since this issue is highly related to the empirical work discussed therein.

As discussed in Section 1.1.4, not every word in a passage of text receives a direct fixation. Rather, short words will often be skipped, as well as highly frequent and predictable words. The skipping of words is explained in the E-Z Reader model by its assumptions about the time course of saccadic programming, and lexical processing. According to E-Z Reader saccades are programmed in two stages with the first and second lasting, on average, 125ms and 25ms, respectively. During the first stage the saccade programme can be cancelled and replaced by a new programme, while in the second stage cancellation is not possible. If the L_1 stage of processing word *n*+1 is completed during the first stage of saccadic programming then the saccade to this word will be cancelled, and a saccade programme to word *n*+2 will be initiated. Consequently, word *n*+1 will be skipped. This sequence of events will occur more regularly in the case of shorter words due to them being closer to the fovea, and thus in high acuity vision, allowing the L_1 stage of lexical processing to be completed more efficiently. Furthermore, high frequency and predictable words will be skipped more often since L_1 is shorter for these words, thus increasing the probability of L_1 completion prior to the first stage of saccadic programming finishing.

The sources of oculomotor error in the execution of saccades observed by McConkie, Kerr, Reddix, and Zola (1988; see Section 1.1.4 and Chapter 4) are also implemented in the E-Z Reader model. The model assumes that saccades are targeted towards the centre of an upcoming word, with the distance between the current point of fixation and this target representing the intended saccade length. In E-Z Reader, a range error is programmed so that for every character that this intended saccade length is either greater than or less than seven characters there will be 0.4 characters of under or overshoot, respectively. Furthermore, there is a parameter for random motor error in the model. This random motor error results in saccades deviating from the centre of a target word, either to the left or the right, and the extent of this potential deviation increases with the length of a saccade. The implementation of these two sources of error allows the model to accurately simulate landing positions within words (Reichle, Rayner, & Pollatsek, 1999).

As discussed in Section 1.2.7 the observation of parafoveal-on-foveal effects are viewed as troublesome for approaches to eye movement control that assume serial lexical processing, such as the E-Z Reader model. It is true that these effects should not typically be observed according to this model. Since the initiation of a saccade programme away from the fixated word is determined by the completion of the L_1 stage of lexical processing, the characteristics of the parafoveal word should not affect the time at which a saccade is programmed away from the foveal word. However, there are several caveats to this. First of all is the fact that the E-Z Reader model includes what is referred to as the pre-attentive visual processing stage. In this stage of processing all visual information available from a display is processed in parallel at the beginning of a fixation, before attention is allocated to a single word. As such, within the serial framework of E-Z Reader it is still possible for highly unusual or salient visual information to be noticed prior to attention being deployed towards this word. Through this pre-attentive visual stage of processing it may be possible for E-Z Reader to explain orthographic parafoveal-on-foveal effects. Reichle, Pollatsek, and Rayner (2006) have proposed that during this stage of visual processing highly unusual letter information in the parafovea may pop out, causing disruption to normal processing. The various studies that have observed orthographic parafoveal-on-foveal effects have created highly unusual visual configurations. Studies showing inhibitory effects tend to use letter sequences which are not encountered during normal reading. Studies showing facilitative effects have relied on repeating the fixated word in the parafovea. During reading the same visual pattern appearing twice in succession is not normal, and it may be that this manipulation is highly salient and thus detected during the pre-attentive stage of visual processing. As such, it may well be that the E-Z Reader model is able to explain such orthographic effects via this mechanism. This is an issue that will be returned to in Chapter 3.

It is also worth addressing lexical parafoveal-on-foveal effects from the perspective of E-Z Reader. There is no way for E-Z Reader to explain these effects if they are in fact real. However, as addressed in Section 1.2.7 there are reasons to doubt that these effects are actually due to the parallel processing of two words, with several alternative explanations existing. One such explanation relies on mislocated fixations, which are a consequence of the saccadic range error. As stated above the E-Z Reader model implements this motor error in its saccadic programming parameters, making this a plausible explanation for lexical parafoveal-on-foveal effects. Furthermore, these effects can also be explained as a consequence of error in an eye-tracker's estimate of the position of the eye. When Reichle and Drieghe (2014) ran simulations using the E-Z Reader model of eye-movement control with added tracker error, the model did indeed predict small lexical parafoveal-on-foveal effects, of the size that have typically been observed in corpus analyses.

1.3.3. The SWIFT Model

The SWIFT model takes a different approach to eye movement control during reading. In this model there is a gradient of attention which is distributed across all words in the perceptual span. In the most recent version of SWIFT (Schad & Engbert, 2012) the spatial extent of the perceptual span is determined by the processing difficulty of the fixated word. The span will typically contain at least word *n* and word *n*+1, and at times may also contain word *n*+2. All of the words within the perceptual span are lexically processed in parallel, with the greatest amount of attentional resources being dedicated to processing the fixated word, and a decreasing amount being deployed to words further into the perceptual span. As such, in SWIFT at least two words tend to be processed simultaneously; this is the main point of contention with E-Z Reader. The processing of these multiple words should occur in independent channels, such that the information which is being processed from one word should not affect the processing of another word. It is through the assumption that multiple words are processed in parallel that the SWIFT model explains the preview effects discussed in Section 1.2.3 and 1.2.4, and the lexical parafoveal-on-foveal effects discussed in section 1.2.7.

According to SWIFT, the time at which we initiate a new saccade program is determined by a random timer, with the saccade being programmed when this timer reaches zero. The rate at which the timer runs down is inhibited by the processing difficulty of the fixated word, such that fixations will be longer on low frequency and unpredictable words (see Section 1.1.3). The where component of the next saccade is determined by an activation field, in which the different words currently being processed compete to be selected as the next saccade target, with the word that has the most activation at the initiation of the saccade program being the target. The level of activation for each word is determined by their current processing status. With regard to word skipping, it is proposed that since the saccade target is determined by the word with the most activation that this word may sometimes be word *n*+2 rather than word *n*+1. As such, readers make a saccade towards word *n*+2 rather than word *n*+1. The rate at which the activation for a potential saccade target rises and falls is determined by the extent to which this word has already been processed; this is influenced by the word's frequency and predictability, as well as how far into the parafovea it extends; this explains the influence of variables such as word length, launch site, predictability, and frequency on skipping probability (see Section 1.1.4).

Lexical parafoveal-on-foveal effects are also explained via the saccade target selection mechanism, and the way in which this depends upon the lexical activation of potential saccade targets. In some cases, the word with the highest activation will be the fixated word; in these cases this word will receive an additional fixation, thus increasing the viewing time for this word. The probability of the fixated word having the highest activation in the perceptual span will be affected by the lexical qualities of the upcoming word (i.e. word *n*+1). On occasions when the lexical qualities of word *n*+1 causes it to have greater activation than word *n*, it will be selected as the saccade target instead of word *n*. In contrast, when the lexical qualities of word *n*+1 leads to it having lower activation than word *n*, it will not be selected as the saccade target instead of word *n*. Consequently, the lexical qualities of word *n*+1 determine whether word *n* receives more than one fixation, and thus the reading time of word *n*. This mechanism within SWIFT explains the observation of lexical parafoveal-on-foveal effects (see Section 1.2.7). It is disputable as to how exactly SWIFT accounts for the existence of orthographic parafoveal-on-foveal effects. Arguably, SWIFT should not predict these effects, due to the fact that it works on the basis of parallel independent channels, rather than parallel interactive channels. As mentioned above, this means that the information encoded as part of each word in the perceptual span should be processed in a separate channel. Each of these separate channels is supposed to be independent of the others, such that the information being processed within one channel should have no influence on the information being processed within another channel. As such, one word either including unusual orthographic information or sharing orthographic information with an adjacent word should not have any effect on reading. Despite this, orthographic parafoveal-on-foveal effects have generally been interpreted as being more in line with the predictions of models such as SWIFT. In Chapter 3 I discuss a framework in which a parallel processing model might predict orthographic parafoveal-on-foveal effects, and present some work testing this framework.

SWIFT makes similar assumptions to E-Z Reader in terms of how saccades are targeted towards words, and the sources of error to which these saccades are susceptible (McConkie, Kerr, Reddix, & Zola 1988; see Section 1.1.4 and Chapter 4). The model assumes that saccades are targeted towards the centre of an upcoming word, with the distance between the current point of fixation and this target representing the intended saccade length. In SWIFT a range error is programmed so that for every character that this intended saccade length is either greater than or less than 5.4 characters there will be 0.41 characters of under or overshoot, respectively.

1.4. Summary and Theoretical Scope

In closing it should be clear that a great deal has been learnt about multiple aspects of eye movement control during reading, including the factors that determine both when and where the eyes move during reading, and how readers integrate information across multiple fixations. Sophisticated computational models — such as E-Z Reader and SWIFT — are able to account for a large amount of the data gathered, and the phenomena that have been observed. Nonetheless, there are still many areas that have either not been fully explored, or that remain controversial. It is my hope that the work presented throughout this thesis will contribute to our knowledge of several areas of eye movement control during reading, and serve to challenge some of the existing assumptions of models of eye movement control. Before proceeding to my first empirical chapter I will briefly outline the questions to be addressed in each study, the relation of these questions to the issues discussed in this review, and my contribution to the published (or to be published) manuscripts.

In Chapter 2 I explore the issue of whether two words can sometimes form a single lexicalised unit, and whether the constituent words of these multi-word units are processed in parallel to a greater extent than two words that do not form a multi-word unit. Specifically, in this chapter I use the boundary paradigm to examine whether word *n*+2 in the parafovea is processed when it is the second constituent of a spaced compound (e.g. *bear* in *teddy bear*), in contrast to prior studies that have failed to observe these effects. This question is important for several reasons; first of all, it is important to understand the nature of the representations that are stored in the lexicon. This question is also relevant to assessing models of eye movement control, since if there are special cases of two words being processed in parallel it raises the question for parallel models of why evidence is not more typically observed of words being processed in parallel; presumably according to these models there should be no special cases. Finally, this chapter relates heavily to the *n*+2 preview effects discussed in Section 1.2.4, and its published edition was actually cited at this point.

The idea that the constituent words of a spaced compound may be processed in parallel as part of a larger multi-word unit was formulated by my co-authors prior to the beginning of my candidature. However, the specific experimental design of this study was formulated by all three authors on the paper. All experimental stimuli were designed by me, and all data acquisition and analysis was primarily performed by me; of course, guidance was provided by both supervisors throughout. Manuscript preparation was a collaborative effort, with me performing the majority of the work and improving the paper on the basis of supervisory feedback on early versions.

In Chapter 3 I address whether orthographic information from one word can influence the processing of another word. This is investigated through the use of the boundary paradigm to present participants with previews in which two letters had either been transposed between word *n*+1 and word *n*+2 (e.g. *hop tan* as a preview of *hot pan*) or in which the same two letters had been replaced by alternative letters (e.g. *hob fan*). Participants gaining a preview benefit from the transposed letter preview relative to the substituted letter preview would be indicative of orthographic information from one word influencing the processing of the other word. If we did find that orthographic information from one word could influence the processing of another, it would be strong evidence in favour of the parallel processing of these two words; thus the study presented in this chapter again holds relevance to the debate between models of eye movement control. The data presented in Chapter 3 are also relevant to the *n*+2 preview effects discussed in Section 1.2.4 and the parafoveal-on-foveal effects discussed in Section 1.2.7. The basic concept of the study presented in this chapter was my own, and I played a primary role in the experimental design. Again, all stimuli design, data acquisition, and data analysis was performed by me, with guidance. The preparation of the manuscript was performed by me, with revisions being made on the basis of supervisory feedback.

Finally, Chapter 4 addresses the extent to which the saccadic targeting system is able to adapt to novel conditions. In this study participants were presented sentences in which all of the words were the same length (e.g. *the sad boy had not had any fun all day*) and the effect of this manipulation upon saccadic targeting was examined with a specific focus on the preferred saccade length (see Section 1.1.4) which is implemented in the two models of eye movement control discussed above. We also examined the effect of this manipulation on word skipping (see Section 1.1.4). While the basic idea of presenting participants with sentences of uniform word length pre-dated the beginning of my candidature, the specific theoretical questions addressed were decided upon as a collaborative effort, with me playing a leading role. Once again, all stimuli design, data acquisition, and data analysis was performed by me, with guidance, and the manuscript was prepared in a similar way to those presented in Chapter 2 and Chapter 3.

Chapter 2

Preview Benefit in English Spaced Compounds

The following Chapter has been published as Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced compounds. *Journal of Experimental Psychology: Learning Memory and Cognition, 40,* 1778-1786.
doi:10.1037/xlm0000013

2.1. Introduction

A well-established finding in research on eye movements during reading is that lexical processing is a primary influence on when to move the eyes (Liversedge & Findlay, 2000; Rayner, 1998; Rayner, Liversedge, White, & Vergilino-Perez, 2003). This is an important assumption of major current models of eye movement control such as E-Z Reader (Reichle, Rayner, & Pollatsek, 2003) and SWIFT (Schad & Engbert, 2012). An assumption of these models that is considerably more controversial is whether words are lexically processed in a serial or parallel fashion during sentence reading. According to E-Z Reader words are lexically processed one at a time in serial order, with processing not beginning on subsequent words until the prior words have been fully identified. In contrast to this, SWIFT proposes that multiple words around the point of fixation are lexically processed simultaneously. On the surface it might seem obvious that the basic lexical unit in English is a word as defined by spaces on either side. However, as we will argue in this paper, this is not necessarily the case, and given the controversy it is essential to establish what can form a lexical unit as this will have direct implications on whether lexical processing during reading occurs in a serial versus parallel fashion.

The question of whether more than one word is lexically processed simultaneously is closely related with how models assume parafoveal information is being processed. Parafoveal processing has mainly been investigated using the boundary paradigm (Rayner, 1975). In this paradigm the preview of a target word is manipulated to be correct or incorrect until the point of fixation crosses an invisible boundary preceding the preview. At this point the preview is replaced with the target word. Reduced fixation times on a word observed after a correct compared to an incorrect preview is known as *preview benefit*. Preview benefit is reliably obtained from one word to the right of fixation (word *n*+1; see Schotter, Angele, & Rayner, 2012), which is predicted by both models. Within SWIFT word *n*+1 is processed at the same time as word *n*. Within E-Z Reader processing of word *n* is completed and the processing of word *n*+1 begins before a saccade is made to word *n*+1. As such, preview benefit effects on word *n*+1 are not controversial.

A considerably more contentious issue is whether preview benefit is obtained from word *n*+2. Reliably gaining a preview benefit this far into the parafovea would be indicative of word *n*+2 being processed simultaneously with preceding words, since ordinarily processing should not begin on word *n*+2 before a saccade is made to word *n*+1. The exception would be when word *n*+1 is very easy to process, allowing it to be completely identified in the parafovea, which would cause the saccade aimed at word *n*+1 to be reprogrammed to word *n*+2, and would result in a short amount of time when attention is on word *n*+2, even though the eyes are still on word *n*. Thus, the false preview of word *n*+2 should not generally be processed often enough for it to have substantial (and reliable) effects upon reading. In contrast SWIFT predicts that word *n*+2 is processed during fixations on word *n* as long as it falls in the perceptual span. The perceptual span in SWIFT is determined by the processing difficulty of the foveal word, and as such, when foveal processing progresses unhindered, SWIFT predicts that *n*+2 preview effects will occur quite often. However, effects of word *n*+2 have only been shown when word *n*+1 was three letters long and highly frequent (Angele & Rayner, 2011; Kliegl, Risse, & Laubrock, 2007; Radach, Inhoff, Glover & Vorstius, 2013), not when it was longer (Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Rayner, Juhasz, & Brown, 2007). Furthermore, the effect has typically been observed during fixations on word *n*+1 rather than *n*+2. Proponents of parallel processing claim that the reason for the effect being restricted to such cases is that a longer word *n*+1 pushes word *n*+2 too far out of the perceptual span. In contrast, advocates of a serial account claim that it is due to word *n*+1 being easily identified because it is so frequent whilst the eyes remain fixated on word *n*. One goal of the current study is to investigate whether an *n*+2 preview benefit can be found in instances where word *n+1* is not as short and highly frequent compared to previous observations of *n+2* preview effects. More specifically, we will examine *n*+2 preview effects when word *n*+2 forms part of a larger linguistic, and possibly lexical, unit with word *n*+1. Were this to be found it would indicate that the constraint on *n*+2 preview benefit is linguistic rather than perceptual.

As mentioned, the idea that the basic lexical unit in English is a word as defined by spaces on either side seems almost trivial, but this is not necessarily the case. Several theories of language posit that the lexicon may contain multi-word units (MWUs). For example, Bybee’s usage-based theory of grammar (2006) proposes that cognitive representations of language are based on experiences with it, and as such commonly occurring MWUs will be lexicalised alongside individual words. Even theories with a distinction between lexical units and the grammar used to build these into sentences allow for the existence of MWUs, such as the Words-and-Rules theory (Pinker, 1999). In this theory the lexicon is comprised of listemes, which are any linguistic units that have to be memorised since they cannot be generated by rules, including MWUs. Thus, while these theories take differing views regarding the representation of language, they both allow for certain MWUs to be lexicalised. Research has shown that MWUs are processed more quickly than non-formulaic language (see Conklin & Schmitt, 2012), suggesting that MWUs may indeed be stored in the lexicon. However, thus far little research has investigated whether MWUs are treated similarly to single words during natural reading.

In the current study a spaced compound is considered to be two frequently co-occurring words which refer to a single concept (e.g *teddy bear*). Given their close relation such MWUs may have a unified lexical entry. That is to say, as well as the lexicon containing separate entries for the words *teddy* and *bear* it may also contain one for *teddy bear*. This possibility is consistent with theories explaining how unspaced compounds (e.g. blackbird) are processed. In dual-route (Pollatsek, Hyönä, & Bertram, 2000) and multiple-route (Kuperman, Schreuder, Bertram & Baayen, 2010) models of compound processing, it is posited that compounds are processed via both a compositional route, in which each constituent is identified separately and then combined to form the compound, and a direct lookup route which accesses a unified lexical entry for the whole compound. Given that the only difference between spaced and unspaced compounds is whether the constituents are spatially separated it is possible that they are processed similarly. Thus, spaced compounds may also have unified lexical entries, which are accessed via a direct lookup route. If unspaced and spaced compounds are identified as single lexical units, then their constituent lexemes should be processed in parallel.

Several studies have investigated whether the constituents of unspaced compounds are processed in serial or parallel, by using the boundary paradigm. Drieghe, Pollatsek, Juhasz, and Rayner (2010) presented invalid previews of the second half of monomorphemic words (e.g. *fountaom* as a preview of *fountain*) and unspaced compounds (e.g. *bathroan* as a preview of *bathroom*) until a boundary in the centre of the word was crossed. Although they observed substantial preview benefits for both types of words, the manipulation only led to significant disruption prior to the eyes crossing the boundary (a parafoveal-on-foveal effect; see Drieghe, 2011) in the case of monomorphemic words. The invalid information at the end of an unspaced compound's second constituent failed to disrupt processing of the first constituent, suggesting that each constituent was processed serially, rather than in parallel.

In contrast, Häikiö, Bertram, and Hyönä (2010) found evidence that the constituent lexemes of unspaced compounds are sometimes processed in parallel. These researchers conducted a boundary study in which an invalid preview of an unspaced Finnish compound's second constituent was presented prior to the eyes moving onto it. In their analysis Häikiö et al. divided their stimuli on the basis of whether the compound was high or low-frequency. Increased fixation durations were observed upon the first constituent when participants were given an invalid preview of the second constituent, but only for the high frequency compounds. Häikiö et al. argued that this demonstrated that the constituent lexemes of high frequency compounds were processed in parallel because high frequency compounds were more likely to be identified via the direct lookup of a unified lexical entry. In contrast, Häikiö et al. suggest that direct lookup processing is less efficient for low than high frequency words, and for this reason, a low frequency compound word is more likely to be identified through the serial processing of its constituent lexemes via a compositional process. Thus, based on Drieghe et al. and Häikiö et al., it appears that the constituents of an unspaced compound can be processed serially or in parallel, depending upon whether a compositional or direct lookup route to identification is more efficient.

Juhasz, Pollatsek, Hyönä, Drieghe, and Rayner (2009) conducted a study using the same boundary manipulation as Drieghe et al. (2010), within spaced compounds. It was found that the preview benefit for the second constituent of a spaced compound was larger than had been observed in prior studies presenting non-compound words with the same types of preview. One possible explanation for this is that the spaced compounds were processed as single lexical units, and as such both constituents were being processed simultaneously. However, not all of Juhasz et al.'s findings were consistent with this explanation. For example, they also observed a level of preview benefit within novel adjective-noun pairs similar to the spaced compound words. Since novel adjective-noun pairs rarely co-occur, then there should not be a unified lexical entry within the mental lexicon corresponding to them. Given this, Juhasz et al. proposed that the second constituents in both adjective-noun pairs and spaced compounds are highly syntactically predictable, thus allowing them to be processed more efficiently in the parafovea. Whilst this alternative explanation may, or may not be correct, for present purposes, our primary focus concerns the possibility that the spaced compounds were processed as single lexical units.

In the current study we positioned the boundary before the first (*n*+1) constituent of a spaced compound, comprised of words *n*+1 and *n*+2. The preview of each constituent was manipulated orthogonally to be either an identity preview, or a non-word (see Figure 2.1). When a saccade crossed the boundary the entirety of the spaced compound was displayed correctly. Importantly, word *n*+1 in the current study was, on average, 5.65 characters long and of low frequency (see Table 2.1), as distinct from the shorter and high frequency words that have been used in other studies. Thus, any effect of the *n*+2 preview might suggest that the two constituent words of spaced compounds are processed as part of a single lexical unit. Furthermore, it would demonstrate that the limiting factor on *n*+2 preview benefit is linguistic rather than perceptual.

The small child gently cuddled his fluffy teddy bear while trying to get to sleep.
The small child gently cuddled his fluffy teddy hocu while trying to get to sleep.
The small child gently cuddled his fluffy fohbg bear while trying to get to sleep.
The small child gently cuddled his fluffy fohbg hocu while trying to get to sleep.

Figure 2.1. An example of the stimuli under different preview conditions. The vertical black line represents the position of the invisible boundary. As the eye crossed this boundary the preview was replaced with the correct version of the spaced compound (e.g. "teddy bear").

We hypothesised that if spaced compounds are processed as single lexical units there should be an *n*+1 preview benefit, and an interaction between the previews of each word, such that there should be an *n*+2 preview benefit only when there is a correct preview of word *n*+1. We reasoned that the first constituent must be present to indicate the compound nature of words *n*+1 and *n*+2. In this way, word *n*+1 licences the processing of the whole of the spaced compound through a direct lookup route.

2.2.1 Method

2.2.2 Participants

Forty-four native English speakers with normal or corrected to normal vision participated. An additional 17 participants were tested but removed from the analysis due to them noticing over three display changes.[1] Participants were rewarded with course credits or £4.50.

2.2.3 Apparatus

Participants' eye movements were tracked using an SR Research Eyelink 1000 system with a sample rate of 1000 hertz. Only movements of the right eye were recorded, although viewing was binocular. Sentences were displayed in black on a grey background, on a single line of a ViewSonic P227f 20 inch CRT monitor, running at a refresh rate of 75 hertz.[2] Viewing distance was 70cm and at this distance 1° of visual angle was occupied by 3.2 characters of monospaced Courier font.

2.2.4 Materials and Design

Forty spaced compounds with a mean transitional probability of 0.42 in the British National Corpus were selected (see Table 2.1). This means that the first constituent appeared as part of the spaced compound 42% of the time within the corpus. These spaced compounds were positioned at least two words from the end of the sentence and were embedded in sentence frames that provided fairly predictive contexts (see Appendix). This was done to make the compounds as easy as possible to process, thus maximising the chances of the spaced compounds being processed while still in the parafovea. A sample item, under the different preview conditions, is shown in Figure 2.1. A cloze task showed that the whole compound was 33% predictable given the sentence up to and including the pre-target word (e.g. *fluffy* in Figure 2.1), and that the second constituent was 97% predictable given the sentence up to and including the first constituent (e.g. *teddy* in Figure

[1] This exclusion rate is not unusual for studies making display changes over multiple words (Angele & Rayner, 2011).
[2] Typically boundary change studies are run at a refresh rate of either 120 or 150 hertz. Unfortunately, due to a technical oversight, the current study was run at 75 hertz. As such, more of the display changes occurred late than is usual for this kind of study. However, we are confident that our data exclusion criteria ensured that the late changes did not contribute to the effects we report.

2.1). Thus, it can be seen that the second constituent of the spaced compounds was highly predictable, with the majority of the predictability deriving from the first constituent.

Using the boundary paradigm (Rayner, 1975) the previews of the first (*n*+1) and second (*n*+2) constituent were orthogonally manipulated to be either an identity preview or a non-word. Hence, the current study used a 2 (*n*+1 preview: Identity vs. Non-word) X 2(*n*+2 preview: Identity vs. Non-word) design. The non-word previews were generated using the same algorithm as Kliegl et al. (2007), which replaced the words with randomly chosen letters, but preserved word shape.

Table 2.1

Length and Frequency Characteristics of the Spaced Compounds, their Constituent Words, and the Pre-Boundary Word. Standard Deviations are Displayed in Parentheses.

Target word	Minimum-maximum length (letters)	Mean length	Mean word frequency per million
Pre-boundary	5-6	5.25 (0.44)	67 (111)
First constituent	4-9	5.65 (0.92)	5 (19)
Second constituent	3-8	4.68 (1.25)	45 (84)
Whole compound	9-14	11.33 (1.54)	3 (15)

Note. Word frequencies were obtained using the British National Corpus.

2.2.5. Procedure

Upon arrival participants were presented with an information sheet and consent form. They were seated in front of the eye-tracker and a head rest was used for stabilisation. Participants were calibrated using a three point horizontal calibration grid, with an acceptance criterion of an average error below 0.33 degrees.

Each trial consisted of a drift check in the middle of the screen followed by a gaze-contingent fixation cross the size of a single character in the position of the first character. If the cross did not trigger or the drift check indicated more than 0.33 degrees of error then the participant was recalibrated. Furthermore, the participant was re-calibrated at regular intervals. When the cross triggered, the sentence appeared. Participants were instructed to read for comprehension, and press a button to move on. There were eight practice sentences. The forty experimental items were mixed in with 69 filler items, 45 of which were part of another gaze-contingent study. On one third of the trials participants were shown a yes/no comprehension question. Responses were made using a game controller.

Across all participants 96% of the questions were answered correctly. The experiment took approximately 30 minutes.

2.3.1. Results

To analyse the data linear mixed-effects models were constructed using the lme4 package (Bates, Maechler, & Bolker, 2012) in R (2013). This type of analysis retains statistical power to a greater extent than ANOVAs in unbalanced designs (see Baayen, 2008), and so is ideal for analysing boundary studies in which trials are often excluded due to early triggers and late display changes. Each preview was treated as a fixed factor with the non-word previews as the baseline, and an interaction term was included. Subjects and items were treated as crossed random factors. Furthermore, two contrasts were programmed to test for *n*+2 preview effects at each level of the *n*+1 preview. The first contrast compared the conditions in which *n*+1 was correct (e.g. *teddy hocu* vs. *teddy bear*), and the second contrast compared the conditions in which the *n*+1 preview was a non-word (e.g. *fohbg hocu* vs. *fohbg bear*).

Prior to analyzing our data we removed fixations above 800ms, and merged any fixations below 80ms with any fixation less than 0.5 degrees of visual angle away, or any fixations below 40ms with any fixation within 1.25 degrees of visual angle. Trials in which the boundary change happened early or participants blinked in a critical region were excluded. Furthermore, trials in which the boundary change completed more than 10ms after fixation onset were excluded, in accordance with Slattery, Angele, and Rayner (2011). These criteria were employed due to a large proportion of late changes (26% of trials). Furthermore, analyses were only conducted on trials where the pre-target word was fixated, in order to ensure that the previews had been seen as words *n*+1 and *n*+2. Altogether these exclusions account for 44% of the data.

Analyses were carried out on words *n* (the pre-boundary word), *n*+1 (the first constituent), *n*+2 (the second constituent), and the whole compound, whereby *n*+1 and *n*+2 constituted a single region. For each interest area five first-pass measures were computed and reading times on a target word were included in the analyses only when participants fixated this word prior to fixating a subsequent word. The first pass measures were first fixation duration (FFD; the duration of the first fixation in a region), gaze duration (GD; the total time between first fixating a region and making a saccade to another region), go-past time (GP; the total time between first fixating a region and making a progressive saccade beyond it), single fixation duration (SFD; the duration of a fixation when it is the only one made on a word), and skipping rate (the proportion of times when a word is not fixated during first pass reading). The means and standard deviations are shown in Table 2.2. The beta values from the models are displayed in Table 2.3. In the case of a significant interaction the beta values from the contrasts are also displayed. The fixation time analyses were carried out on log-transformed data to increase normality. The skipping data were analysed using logistic models due to the binary nature of the variable.

2.3.2. Word *n*

There were no significant differences between conditions on word *n*.

Table 2.2

Fixation Time Measures (in Milliseconds) and Fixation Probability for all Target Regions. Standard Deviations are Presented in Parentheses.

Preview	Word *n*	Word *n*+1	Word *n*+2	Whole compound
		First fixation duration		
Both identity	232(69)	230(69)	197(73)	-
n+2 non-word	230(78)	246(74)	206(74)	-
n+1 non-word	234(72)	286(91)	198(74)	-
Both non-word	235(72)	267(89)	196(60)	-
		Gaze duration		
Both identity	247(100)	241(76)	211(90)	354(143)
n+2 non-word	239(96)	268(83)	220(86)	393(161)
n+1 non-word	248(85)	325(97)	207(83)	437(157)
Both non-word	249(92)	323(103)	206(79)	464(157)
		Go-past time		
Both identity	253(110)	249(82)	211(90)	375(151)
n+2 non-word	247(97)	276(92)	226(93)	408(172)
n+1 non-word	253(94)	335(101)	215(93)	468(163)
Both non-word	255(97)	334(110)	210(82)	484(157)
		Single fixation duration		
Both identity	231(69)	235(70)	193(67)	226(72)
n+2 non-word	232(80)	256(75)	207(74)	257(78)
n+1 non-word	238(74)	305(87)	196(76)	296(90)
Both non-word	237(73)	295(87)	195(60)	313(74)
		Skipping probability		
Both identity	-	.09(.29)	.33(.47)	.01(.11)
n+2 non-word	-	.18(.39)	.26(.44)	.01(.11)
n+1 non-word	-	.07(.26)	.33(.47)	.00(.06)
Both non-word	-	.07(.26)	.25(.43)	.00(.07)

Note. Skipping data for word n is not available due to its exclusion in the data cleaning procedure. First fixation duration data for the whole compound is not presented for reasons discussed in the text.

2.3.3. Word *n*+1

There was a significant effect of *n*+1 preview type in all measures, such that fixation times were shorter on word *n*+1 when participants had received an identity preview rather than a non-word preview, thus replicating Rayner (1975). More interestingly, there was an interaction between the two previews across all fixation time measures. The planned contrasts showed that in the case of all measures this was due to a significant *n*+2 preview benefit when *n*+1 was available but not when *n*+1 was unavailable. The one exception to this was in FFD, where there was a significant main effect of the *n*+2 preview type in both contrasts. The effects in both contrasts went in opposite directions depending upon the *n*+1 preview, such that FFDs on word *n+1* were shorter for an identity *n*+2 preview when *n*+1 was an identity preview but longer when word *n*+1 was a non-word preview. This effect disappeared in later measures. We suspect this pattern to be due to an increased number of second fixations when the whole compound was disrupted (n=55) than when only word n+1 was disrupted (n=38), given that first fixations tend to be shorter when there is a second fixation (Rayner, Sereno, & Raney, 1996).

Finally, there was a marginal interaction on skipping probabilities, with a significant *n*+2 effect such that an invalid *n*+2 preview led to increased skipping of *n*+1 when *n*+1 was available, but not when *n*+1 was unavailable. This effect suggests that the disrupted *n*+2 attracted attention, but only when word *n*+1 was undisrupted.

In summary, we found both an *n*+1 preview benefit and strong evidence for an *n*+2 preview benefit when the preview of the first constituent was correct.[3]

[3] To ensure that the restrictions on our data set did not contribute to any statistical effects, we also conducted further analyses on our data set, using both more stringent, and more liberal, exclusion criteria. In one analysis we applied the existing exclusion criteria, and also excluded trials in which the boundary change was even 1 ms late (leaving 44% of the data). In this data set the slower refresh rate could not have influenced our results. The overall pattern was very similar, and critically, the pattern of effects was identical in the n+1 target region. Specifically, the effect in the n+1 target region for this analysis showed that there was an n+2 preview benefit given a valid n+1 preview of 16, 22, 24, and 15 ms for FFD, GD, GP, and SFD, respectively. Furthermore, we conducted an analysis which included trials in which the pre-target word was skipped or the boundary was triggered during a saccade that landed on word n (leaving 76% of the data). We originally excluded these trials due to the target words not being previewed, and therefore, the addition of these trials could have weakened our effect. However, the pattern in the data remained significant and the same, with the preview benefit amounting to 15, 19, 16, and 17 ms in the various measures. Given the similarity of these different

2.3.4. Word *n*+2

There were no significant effects in this region.

Table 2.3

Fixed Effect Estimate from the LME Models for all Measures across All Regions

Factor	First fixation duration	Gaze duration	Go-past time	Single fixation duration	Skipping probability
			Word *n*		
n+1 preview	-0.02	-0.04	-0.02	-0.02	-
n+2 preview	0.00	0.00	0.00	0.01	-
Interaction	0.01	0.02	0.01	-0.02	-
			Word *n*+1		
n+1 preview	-0.07*	-0.19***	-0.19***	-0.14***	1.19***
n+2 preview	0.08**	0.01	0.01	0.03	0.00
Interaction	-0.14***	-0.12**	-0.11**	-0.13**	-0.89+
First contrast[a]	-0.08**	-0.13***	-0.13***	-0.13***	-1.04**
Second contrast[b]	0.08**	0.02	0.02	0.05	-0.00
			Word *n*+2		
n+1 preview	0.04	0.06	0.07	0.05	0.03
n+2 preview	-0.01	0.00	0.00	-0.01	0.37
Interaction	-0.04	-0.04	-0.08	-0.05	0.01
			Whole compound		
n+1 preview	-	-0.17***	-0.19***	-0.22**	1.23
n+2 preview	-	-0.05	-0.03	-0.08	-0.20
Interaction	-	-0.07	-0.07	-0.05	0.04

[a]Refers to the comparison between the two conditions in which an identity preview of word *n*+1 was given. [b]Refers to the comparison between the two conditions in which there was a non-word preview of word *n*+1.

*p<.05

**p<.01

***p<.001

+p<.10

2.3.5. Whole compound

FFD was not examined in this region since it would mainly consist of the same data as for word *n*+1. While there was a significant effect of the *n*+1 preview across all fixation time measures, the effects of the *n*+2 preview and the interaction failed to reach significance in any measure.

analyses it seems reasonable to assume that our effect is robust, and does not arise as an artefact of the refresh rate or our exclusion criteria.

2.4. Discussion

The current study tested whether *n*+2 preview effects could be observed when *n*+1 and *n*+2 were constituents of a spaced compound. The existence of *n*+2 preview benefit has been controversial with findings limited to experiments using a short and highly frequent word *n*+1. The current experiment demonstrates reliable and sizeable *n*+2 preview effects when *n*+1 and *n*+2 constitute a spaced compound. Because *n*+1 was longer and lower frequency than in previous experiments showing *n*+2 preview effects, this experiment convincingly demonstrates that previous failures to find *n*+2 effects were not necessarily due to *n*+2 being too far into the parafovea. It appears that when lexical processing of word *n*+1 licenses parafoveal processing of word *n*+2, a parafoveal preview benefit of word *n*+2 can be observed.

It was hypothesised that if spaced compounds are processed as lexical units then an interaction would be observed, such that there would be an *n*+2 preview benefit, but only when the first constituent of the spaced compound was available. This pattern of effects is exactly what was found in fixation times on word *n*+1. This preview benefit was 16ms in first fixation duration, 21ms in single fixation durations, and 27ms in gaze duration and go-past time. This suggests that processing of the second constituent of the spaced compound occurred while it was two words to the right of fixation, but only if this was licensed by the first constituent indicating the compound nature of the stimuli. It is our contention that this is due to the two words having a unified lexical entry, which is identified as a single unit through a direct lookup route after an initial period of compositional processing.

A full specification of how processing may be extended further into the parafovea due to multiple words forming a single lexical unit is beyond the scope of the current article. However, one way in which we envision this occurring is through feed-down activation in the context of an interactive-activation framework (McClelland & Rumelheart, 1981). In this approach processing would begin on the first constituent of a spaced compound in the parafovea, causing excitation of the lexical entries for both the individual constituent, and the spaced compound it is a part of. Both these lexical entries would become activated and then feed activation back down to the letter level of the lexical processing system. The activation that was fed down from the lexical entry of the spaced compound would activate letters associated with both the first and the second lexeme of the compound. This, along with orthographic information about the second lexeme extracted from the parafovea, would boost the activation of these letters leading to facilitated identification of the entry associated with the spaced compound at the word level. Future work involving both formal modelling and further empirical investigation of lexical identification of spaced compounds is required to fully develop and evaluate this explanation.

The fact that the *n*+2 preview effect appeared during fixations on word *n*+1 is consistent with prior *n*+2 boundary studies. Debate exists as to whether this is typically due to fixations targeted towards word *n*+2 undershooting and landing on word *n*+1, or a delayed cost associated with processing the false preview of *n*+2 from word *n*, which itself does not appear in the eye movement record until fixations on word *n*+1 (see Risse & Kliegl, 2012). It is our contention that an alternative explanation is more plausible for the current study. If words *n*+1 and *n*+2 are processed as parts of a larger MWU then processing of word *n*+1 occurs simply as part of processing associated with the larger MWU (and therefore simultaneously with word *n*+2). Given that on ninety per cent of trials word *n*+1 was fixated prior to word *n*+2, it is hardly surprising that this is the region where we observed effects of the display change.

The lexicon containing units larger than single words is in line with several theoretical accounts (e.g. Bybee, 2006; Pinker, 1998). Within these theories spaced compounds are only one subset of MWUs that may be lexically represented, with other candidates including idioms, clichés, collocations, binomial word pairs (e.g. *bride and groom,* see Siyanova-Chanturia, Conklin, & van Heuven, 2011) and other common phrases. It is important to note that these theories vary with regard to which MWUs are lexicalised and which are not. According to the Words-and-Rules theory only MWUs that cannot be generated out of smaller lexical units via rules should have lexical entries (Pinker, 1998). In contrast, usage-based theories propose that all commonly occurring MWUs should be lexicalised (Bybee, 2006). Given that the current study strongly suggests that one type of MWU is indeed lexicalised, it is important for future research to establish what other kinds of MWU are and are not lexicalised, and what the criteria for lexicalisation are. By establishing such criteria it will be possible to resolve some of the points of dispute between current theories.

While we have framed our results in terms of lexicalised MWUs some researchers may consider that our findings arose solely as a consequence of how predictable word *n*+2 was. Indeed, current models of eye movement control state that lexical identification is linked to predictability. Furthermore, much like in Juhasz et al.'s (2009) study, there would have been a high degree of syntactic predictability for the second constituent of the spaced compounds given the first constituent. However, while it has been found that predictability influences the degree of preview benefit from word *n*+2 (Radach et al., 2013), we believe there are good reasons why predictability is a less feasible account of our findings than our targets forming lexicalised MWUs. In the current study the predictability of word *n*+2 arose predominantly due to word *n*+1 rather than the preceding context. Recall, word *n*+2 was only 33% predictable from the sentence up to the pre-target word, and became 97% predictable given word *n*+1. As such, for the high predictability of word *n*+2 to have driven our effect it would have been necessary for word *n*+1 to be identified and integrated into the sentential context during fixations on word *n*. This is true from the perspective of both serial and parallel models. It is unlikely that this occurred reliably enough to have driven our effect, given that word *n*+1 was both long and low frequency. The fact that prior studies have failed to find *n*+2 preview effects when *n*+1 is long supports this idea. In Radach et al.'s study, in which *n*+2 effects were obtained, word *n*+1 was always "the", making it highly likely to be identified and integrated during fixations on word *n*. As such, it is unlikely that Radach et al.'s and our findings arose for the same reasons. Rather, the more plausible explanation for our finding is that word *n*+1 was parafoveally processed to an extent that the compound nature of the two upcoming words became clear during fixations on word *n,* and thus, processing of the second constituent was licensed as part of a lexicalised MWU. It is important to note that our explanation does not reject the role of predictability, and that these two explanations are not mutually exclusive. It is possible that the high predictability provided by the first constituent of the spaced compounds contributes to the licensing process. Where our position diverges from a standard predictability account is that we propose that predictability merely contributes to early processing of *n*+2 as part of a lexicalised MWU, as opposed to it simply being processed more efficiently once *n*+1 has been fully identified.

Arguably, our findings may pose issues for models of eye movement control such as E-Z Reader (Reichle, Rayner, & Pollatsek, 2003) and SWIFT (Schad & Engbert, 2012). Currently these models do not take into account that some lexical entries could be composed of multiple words and may be processed as single units. Given that some analyses suggest that about 50% of written discourse may consist of MWUs (Erman & Warren, 2000) this might be something that should be incorporated into the models. Within E-Z Reader word identification proceeds serially and sequentially, with lexical processing only beginning on a word after all preceding words have been fully identified. This was clearly not the case in the current study, with processing of the second constituent of a spaced compound occurring whilst the word before the compound was still fixated. While this finding may seem problematic for the idea of serial processing, we believe that it is not, due to our position that spaced compounds may be processed as single lexicalised MWUs. Under this viewpoint lexical processing would have still only encompassed one lexical unit at a time, with this lexical unit consisting of two letter strings separated by a space.

It is clearly not an issue for parallel models that two parafoveal words are processed in parallel,[4] with the models predicting such effects. However, our findings do become problematic when considered in the wider context of prior studies investigating *n*+2 preview benefits. Proponents of parallel models argue that the reason such effects are not found when word *n*+1 is longer is because under such circumstances *n*+2 is no longer in the perceptual span (e.g. Kliegl, Risse, & Laubrock, 2007). The current study brings this suggestion into question, since the second constituent of our spaced compounds was quite far to the right of fixation yet still produced preview effects. Thus, it seems more likely that the extent to which information is processed parafoveally is influenced by linguistic factors, and in the case of the current study whether the parafoveal words form a single lexical unit.

In closing, the current study has extended and supported prior work on the lexical representation of spaced compounds by showing that they may be processed as single units in the parafovea. We believe this to be one of the strongest pieces of evidence thus far in favour of MWUs having unified lexical entries. Given how prevalent such units may be it is important to gain a clearer understanding of how they are processed during natural reading and what it is that causes certain MWUs to become lexicalised. Thus, it seems likely that the degree to which words are processed in the parafovea is not only constrained by perceptual limitations, but also the linguistic characteristics of those words. Specifically, in the case of the current study, whether the words form a single lexical unit.

[4] While our main finding is not problematic for a parallel model, it is interesting to note the null effect of our preview manipulation on reading word n. This constitutes a failure to replicate controversial orthographic parafoveal-on-foveal effects. This was true even in the case of when both words n+1 and n+2 were non-words compared to when both were correct. The extent of the illegal information here is greater than in many prior studies, and yet there was no effect. As such our data strongly contradicts the idea of orthographic parafoveal-on-foveal effects, which are important for parallel accounts of processing.

Chapter 3

Is Orthographic Information from Multiple Parafoveal Words Processed in Parallel?

At the time of thesis submission the following chapter is under review as Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2016). *Is orthographic information from multiple parafoveal words processed in parallel: An eye-tracking study*. Manuscript submitted for publication.

3.1.1. Introduction

During reading visual information is simultaneously available from the fixated word, as well as several spatially adjacent words. As such, multiple words are available for processing during a single fixation. However, the manner in which readers make use of the parafoveal information is disputed, both in terms of the spatial extent of parafoveal processing, and the time course with which this processing takes place. According to the serial approach (Reichle, Liversedge, Pollatsek, & Rayner, 2009; Reichle, Rayner, & Pollatsek, 2003) only one word is processed at a time during reading, with the processing of subsequent words not beginning until the fixated word has been fully identified. In contrast, the parallel approach (Engbert, Nuthmann, Richter, & Kliegl, 2005; Schad & Engbert, 2012) proposes that several words around the point of fixation are processed simultaneously. These two approaches make differing predictions with regard to the number of words processed during a single fixation, and the time course of such processing.

One issue which is often not addressed in this debate is how the parallel lexical processing of multiple words may actually operate. Reichle et al. (2009) posed several problems that may arise during reading as a consequence of processing multiple words in parallel. One such issue relates to how orthographic information from multiple words would accurately lead to the activation and identification of each separate lexical item, rather than noisy activation distributed across multiple lexical entries and corresponding precisely to neither word. Reichle et al. argued that there are two levels of processing at which this could become problematic. The first is at the point of encoding the words, when focussed attention is likely to be necessary to bind the features of a visual object together;

in the case of words, the constituent letters. Presumably, if attention was divided across multiple words then the letters of the words may not be adequately bound, and as such the activation caused by each letter may not be specific to a particular word. Furthermore, Reichle et al. made the point that most models of word identification involve the orthographic units of a word feeding activation into the lexicon, with these orthographic units causing activation at higher levels of representation. If orthographic units from more than one visually presented word caused activation in the lexicon simultaneously it is unclear what sort of mechanism would prevent the orthographic units of one word interfering with the processing of the other, and vice versa. An implication of both issues (the need to focus attention to bind the constituent letters of a word, and what happens if orthographic units from multiple words cause lexical activation at the same time) is that if two words were lexically processed in parallel then we may expect the orthographic units of one word to influence the processing of another. If evidence of this occurring was to be found it would constitute very strong support for the parallel processing of multiple words. In the current study we explore these issues by investigating whether orthographic information available from one parafoveal word affects the lexical processing of an adjacent parafoveal word.

Before proceeding to outline our own experiments, it is important to discuss how different approaches to lexical processing have been formalized in models of eye movement control, and how these approaches account for prior findings relating to the processing of parafoveal information. It is especially important to understand the parallel approach, with the current study aiming to find evidence in favour of this, and it is also necessary to understand the serial approach in terms of how it is able to explain prior findings that on the surface may already seem to constitute evidence in favour of parallelism.

One of the models that embodies the parallel processing approach is the SWIFT model (Engbert, Nuthmann, Richter, & Kliegl, 2005; Schad & Engbert, 2012). In SWIFT, a perceptual span around the point of fixation is determined by the processing load of the fixated word, with a less cognitively demanding word resulting in a broader perceptual span. The perceptual span will typically contain multiple words, with all of these words being processed in parallel. The serial position is implemented in the E-Z Reader model (Reichle, Pollatsek, & Rayner, 2006; Reichle, Rayner, & Pollatsek, 2003; Reichle, Warren, & McConnell, 2009). In this model lexical processing proceeds in two separate stages (L_1 and L_2). In the first stage (L_1) the orthographic familiarity of a word is assessed; in the second stage (L_2) the word is lexically accessed and its meaning is retrieved. Upon the completion of L_1 the oculomotor control system initiates the programming of a saccade to the next word (word *n*+1), while processing continues upon the fixated word (word *n*). Thus, the programming of a saccade and the later stages of lexical processing (L_2) occur concurrently. Upon word *n* being fully identified attention shifts to word *n*+1, and this word is processed in the parafovea until the saccade program is completed, at which point word *n*+1 is directly fixated. A further important feature of the E-Z Reader model to the current study is the pre-attentive visual processing stage, in which information from the entire visual field is passed to the brain in parallel, prior to attention being allocated to a single word. Due to these assumptions the two models make differing predictions with regard to the processing of parafoveal information.

The processing of parafoveal information during reading has primarily been investigated using the boundary paradigm (Rayner, 1975). In the boundary paradigm an invisible boundary is placed before a target word, and prior to the eye crossing this boundary there is either a correct or incorrect preview of the target word, which changes to the target word as the boundary is crossed (see Figure 3.1). This technique can be used to investigate a number of issues relating to the processing of parafoveal information, including the extent to which the characteristics of an upcoming word influence the processing of the fixated word, and the extent to which a word is processed prior to direct fixation. Both of these issues are relevant to the current study, and we will explore them in turn.

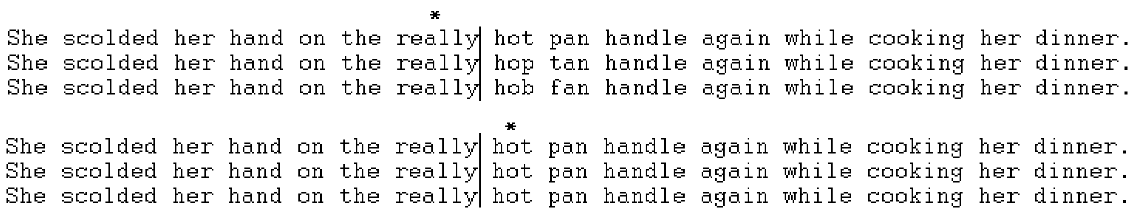


Figure 3.1. An illustration of the stimuli and the boundary paradigm. The vertical black line represents the invisible boundary, and the asterisk the point of fixation. Prior to the eye crossing the boundary participants are presented with either an identity preview (hot pan), a preview in which two letters have been transposed (hop tan), or a preview in which the same two letters have been substituted (hob fan). After the eye crosses the boundary, the previews become the target words in all conditions.

3.1.2. Parafoveal-on-Foveal Effects

The issue of whether the orthographic information which is available from one word in the perceptual span can influence the processing of a word that appears earlier in the perceptual span has previously been addressed in experiments investigating parafoveal-on-foveal effects (for a review, see Drieghe, 2011). Parafoveal-on-foveal effects occur when the characteristics of a parafoveal word (typically word *n*+1) affect fixation durations on word *n*. Such effects have been observed in relation to the orthography of an upcoming word, as might be expected in a parallel processing account according to the arguments of Reichle et al. (2009). Several studies have shown that orthographically illegal or unusual information at the start of a parafoveal word can influence fixation times on the foveal word (Drieghe, Rayner & Pollatsek, 2008; Inhoff, Starr, & Shindler, 2000). In these studies participants fixated a target word for longer when the parafoveal word was manipulated to be orthographically illegal (e.g. *pvxforming*) using the boundary paradigm, than when more typical information (e.g. *performing*) was presented. Furthermore, White (2008) observed inflated fixation times on a word when it was followed by an orthographically unfamiliar (e.g. *crypt*) relative to an orthographically familiar word (e.g. *adder*). It should be noted that these orthographic parafoveal-on-foveal effects are not always observed. For example, White and Liversedge (2006) found that participants did not fixate a target word for any longer when the following word was misspelt (e.g. *pwformer*) than when it was spelt correctly (e.g. *performer*). Nonetheless, it remains worthwhile considering whether these effects may be evidence for the parallel processing of multiple words within the perceptual span.

In addition to these detrimental effects of illegal orthographic information, research has demonstrated facilitative effects of parafoveal orthographic information on foveal word processing. Angele, Tran, and Rayner (2013) examined fixation durations on a target word (e.g. *news*) while using the boundary paradigm to manipulate the parafoveal word to be a repetition of the target (e.g. *news*), an orthographically similar non-word (e.g. *niws*), semantically related (e.g. *tale*), an orthographically dissimilar non-word (e.g. *tule*), or the post-boundary word itself (e.g. *once*). They found that readers fixated the word prior to the invisible boundary for less time when it shared orthographic information with the parafoveal preview than in the other conditions (see also Dare & Shillcock, 2013; Inhoff, Radach, Starr, & Greenberg, 2000). While there were facilitative parafoveal-on-foveal effects from this repetition, the incorrect preview did lead to inflated fixation durations on word *n*+1, as is typical in studies using the boundary paradigm. Similarly, isolated word recognition studies have demonstrated facilitative effects of repeating the letters of a target word to either side (e.g. *RO ROCK CK*) on lexical decision times relative to a condition in which the target word was flanked by unrelated letters (e.g. *ST ROCK EN*) (Dare & Shillcock, 2013; Grainger, Mathôt, & Vitu, 2014).

These studies seem to suggest that it is indeed possible for orthographic information from one word to influence the processing of another, in a manner similar to what would be predicted for parallel processing according to the arguments of Reichle et al. (2009). Specifically, the facilitative effects of overlapping orthographic information documented above could be due to two sets of identical orthographic information being fed into the mental lexicon at once, leading to the increased activation of that word's orthographic representation, and thus the rapid identification of that word. If this were the locus of this effect, it would certainly suggest a level of parallel lexical processing. Presumably within this framework the inhibitory effects of illegal parafoveal orthographic information would be due to the unusual information being fed into the lexicon having an adverse effect upon attentional resources, and thus the processing of word *n*.

However, there is an alternative explanation for these effects which is in line with a serial processing model, rather than the parallel processing of multiple words. As stated earlier the E-Z Reader model includes a pre-attentive visual processing stage, in which all available visual information is processed in parallel. It has been proposed (Reichle, Pollatsek, & Rayner, 2006) that during this early stage of processing, unusual letter information may pop out due to its visual saliency, and interfere with normal processing. The manipulation of parafoveal information in all of the above studies involved the manipulation of visual (orthographic) configurations. In the case of the studies showing inhibitory parafoveal-on-foveal effects the letter sequences were highly unusual, and thus salient. In the studies showing facilitatory parafoveal-on-foveal effects the repetition of the same visual pattern in close proximity (a very rare occurrence during reading) was highly visually salient. Thus, current observations of orthographic information from one word influencing the processing of another are not necessarily evidence for multiple word forms feeding activation into the lexicon in parallel; rather, these effects may simply be due to the visually unusual nature of the manipulations used in these studies.[5] As such, a stronger test

[5] It should also be noted that there are various other explanations for parafoveal-on-foveal effects in a serial model. These explanations typically view the effects as being artefacts of other phenomena such as mislocated fixations (Drieghe et al., 2008) and eye-tracker error (Reichle & Drieghe, 2014). While

is required of the possibility that orthographic information from multiple words within the perceptual span is fed into the lexicon in parallel, and that the information from one word is able to influence the processing of another word. Furthermore, it is necessary for this test to avoid confounding this orthographic manipulation with increased visual saliency. In the current paper we present such a test across two experiments, by transposing an orthographic unit (i.e. letter) from one word in the parafovea with an orthographic unit from another word in the parafovea. Under the assumption that both of these parafoveal words are processed prior to fixation, and in parallel, we should observe an effect of this manipulation relative to a condition in which these orthographic units are simply replaced by unrelated ones. Thus, it is necessary to establish how prior research (and different theoretical approaches) suggests these two parafoveal words should be processed.

3.1.3. Preview Benefits during Reading

The boundary paradigm has more typically been used to examine the extent to which a word is processed prior to direct fixation, in terms of the way in which having a correct relative to incorrect preview of an upcoming target word will influence fixation durations on that word. Typically, participants will spend less time fixated on the target word given a correct rather than incorrect preview, an effect known as preview benefit. In addition to a preview benefit being observed for a correct relative to incorrect preview, it is also reliably observed for previews that share an abstract characteristic such as orthographic, phonological, or semantic information with the target word (see Cutter, Drieghe, & Liversedge, 2015 for a recent review). Preview benefit effects demonstrate that the parafoveal word has been pre-processed, and has activated a candidate set of lexical items, the members of which are congruent with the available parafoveal information. It has long been uncontroversial that when the previewed word appears directly after the boundary (i.e. word *n*+1) participants reliably gain a preview benefit. Both serial and parallel approaches can account for this finding. According to E-Z Reader attention has shifted to word *n*+1 prior to a saccade being made, while in SWIFT word *n*+1 typically falls in the perceptual span and is thus processed prior to direct fixation.

More recent research has investigated whether a preview benefit is obtained when there is an additional word between the boundary and the target word (i.e. the preview

---

these explanations are able to explain the inhibitory effects discussed above, they cannot account for the facilitative effects.

manipulation occurs for word *n*+2). This question is viewed as a further testing ground on which to distinguish between serial and parallel approaches. In the E-Z Reader model word *n*+1 should not typically be identified prior to a saccade being made across the boundary, and thus a substantive preview of word *n*+2 should usually not be obtained. In SWIFT word *n*+2 should regularly fall into the perceptual span, and consequently a preview of this word should be obtained regularly from word *n*. As such, large and prevalent effects of an *n*+2 preview would be more in line with a parallel processing account than a serial processing account. However, research has shown that readers only obtain small preview benefits from word *n*+2, and only under restricted circumstances, such as when word *n*+1 is only three letters long (Angele & Rayner, 2011; Kliegl, Risse, & Laubrock, 2007; Radach, Inhoff, Glover & Vorstius, 2013; Risse & Kliegl, 2012; 2014), or when it forms a larger lexical unit with word *n*+1 (Cutter, Drieghe, & Liversedge, 2014). Effects have not been observed when word *n*+1 was longer than three letters (Angele, Slattery, Yang, Kliegl, & Rayner, 2008; Rayner, Juhasz, & Brown, 2007).

This pattern of results is actually compatible with both models. Within E-Z Reader a short word *n*+1 is easy to identify, allowing attention to shift to word *n*+2 prior to a saccade being made across the boundary (see Schotter, Reichle, & Rayner, 2014 for a simulation). Within SWIFT a longer word *n*+1 causes word *n*+2 to be further away from fixation, and thus less likely to fall into the perceptual span. Furthermore, even when it does fall within the perceptual span it is in lower acuity vision, and thus processed less efficiently. As such, efforts to distinguish between serial and parallel accounts of lexical processing through the investigation of *n*+2 preview effects have thus far been inconclusive, with it being unclear whether effects are due to the parallel processing of word *n*+1 and word *n*+2, or the serial processing of these words with the rapid identification of word *n*+1 from word *n*.

3.1.4. The Current Study

In summary, existing research on both *n*+2 preview benefits and orthographic parafoveal-on-foveal effects is unable to adequately discriminate between serial and parallel processing accounts. With regard to orthographic parafoveal-on-foveal effects it may be that unusual orthographic information in the parafovea is detected at a pre-attentive visual level, rather than that orthographic information from multiple words is being encoded and fed into the lexicon in parallel. In the case of *n*+2 preview effects it is possible that word *n*+1 is fully identified prior to a saccade being made from word *n*, thus leading to the processing of word *n*+2 as a result of a serial attention shift, as opposed to both word *n*+1 and *n*+2 being lexically processed at the same time. In the current paper we investigate both issues simultaneously, by using the boundary paradigm to examine whether orthographic information extracted from word *n*+1 while it is in the parafovea influences the processing of word *n*+2, and vice versa.

In both of our experiments participants received an identity preview of word *n*+1 and word *n*+2 (e.g. *hot pan*), a preview in which a letter had been transposed between these words (e.g. *hop tan*), or a preview in which the same two letters were substituted (e.g. *hob fan*). In Experiment 1 this manipulation always involved the final letter of word *n*+1 and the initial letter of word *n*+2, whereas in Experiment 2 the manipulation was always applied to the same letter position in each word (e.g. *pit hop* and *fit cop* as previews for *hit pop*). We hypothesized that if word *n*+1 and *n*+2 were processed in parallel we may observe a preview benefit for the condition in which letters had been transposed between words rather than substituted. We made this prediction on the basis that parallel processing accounts assume that *n*+2 preview effects are due to the processing of word *n*, *n*+1, and *n*+2 at the same time. Furthermore, the simplest explanation for facilitative orthographic parafoveal-on-foveal effects in a parallel approach is that they are due to the extraction of orthographic information from two spatially adjacent words in parallel, with both sets of orthographic information being fed into the lexicon at the same time. Following the arguments of Reichle et al. (2009), the orthographic information from one word would then influence the processing of the other word. If both of these related contentions are correct, it seems reasonable to propose that participants may gain a benefit from having correct letter identity information from both words in the parafovea, even if that orthographic information is not position specific. Furthermore, unlike studies that have previously observed facilitative orthographic parafoveal-on-foveal effects, our manipulation did not involve the use of visually salient stimulus patterns in the parafovea, and consequently any effects that might be observed could not be due to pre-attentive visual processing. Thus, if we did observe a benefit of the transposition condition relative to the substitution condition, it could only be due to processing of word *n*+1 and word *n*+2 occurring in parallel. In contrast, a serial model would predict no differences between these two conditions, since the processing of information from word *n*+1 and word *n*+2 should be independent. Rather, a serial model would only predict a preview benefit of the identity condition relative to the two alternative preview types. It should, however, be noted that a null effect of our transposition relative to substitution preview would not necessarily provide evidence against parallel lexical processing, so much as a lack of support for this specific position.

3.2.1. Experiment 1

Our parafoveal preview manipulation in Experiment 1 involved making a transposition between the final letter of word *n*+1 and the initial letter of word *n*+2, such that the transposed letter preview for the target words *hot pan* would be *hop tan*, with a substitution preview of *hob fan*. Word *n*+1 was always three letters long, while word *n*+2 could either be three or four letters long. As such, our manipulation was always made between the third and fifth character beyond the end of word *n* (the space after word *n+1* being the fourth character). Furthermore, word *n*+1 was the same length as in prior studies demonstrating *n*+2 preview effects (e.g. Risse, Kliegl, & Laubrock, 2007). The previews always formed two new words, as opposed to non-words. This approach was taken due to the fact that prior research has shown that people do not process word *n*+2 when word *n*+1 is a non-word (Angele & Rayner, 2011). As such, every precaution was taken to ensure that word *n*+2 should typically have been processed from word *n*, regardless of whether this was due to the parallel processing of multiple words or a serial attention shift towards it.

There are obvious parallels of our manipulation in Experiment 1 with the transposed letter effect in word identification (see Frost, 2015 for a recent review and discussion). The transposed letter effect refers to a phenomenon whereby either a prime or parafoveal preview of a target word (e.g. *judge*) in which two letters are transposed (e.g *jugde*) leads to faster recognition times for the target word than a preview in which the same letters have been substituted (e.g. *jupte*). While we did not set out to specifically investigate the transposed letter effect so much as the processing of multiple words in the parafovea, there are some findings from this literature that are briefly worth discussing in relation to our own study.[6] First, several prior studies have shown that readers gain a greater parafoveal preview benefit from a transposed letter preview relative to a substituted letter preview (Johnson, 2007; Johnson & Dunne, 2012; Johnson, Perea, & Rayner, 2007;

[6] There is an extensive literature examining the locus of this effect in terms of visual word identification. We do not explore this literature in the current article due to our primary concern being the way in which the orthographic information from one word influences the processing of an adjacent word.

Pagán, Blythe, & Liversedge, 2016). Of particular relevance to the current experiment is that Johnson (2007) observed these effects when applying the manipulation to the third and fifth letter in a parafoveal preview of a target word, such that the word *flower* would receive shorter fixations given the parafoveal preview *flewor* as opposed to *flawur*. In the current experiment our manipulation was always made between these same two letter positions in the parafovea, albeit across a space, rather than within a word. Thus, within a situation in which the parallel processing of letters is uncontroversial prior research suggests that correct letter identity information in the parafovea can provide a preview benefit, even when these letters have been transposed across two character spaces. Clearly, Johnson's study suggests that readers are generally capable of detecting a manipulation of the magnitude used in the current experiment. Thus, any failure to observe an effect in the current experiment would most likely be a result of the way in which people process parafoveal information across two words, relative to within one word.

3.2.2.1. Method.

3.2.2.2. Participants.

Forty-eight native English speakers with normal or corrected to normal vision participated. An additional 13 participants were tested but removed from the analysis due to noticing more than five display changes. Participants were rewarded with research credits or £4.50.

3.2.2.3. Apparatus.

An SR Research Eyelink 1000 system with a sampling rate of 1000 hertz was used to track participants' eye movements. Sentences were displayed on a single line of a ViewSonic P227f 20 inch CRT monitor with a grey background, running at a refresh rate of 75 hertz.[7] Viewing distance was 70cm, with 1° of visual angle containing 3.2 characters of monospaced courier font.

[7] Typically boundary change studies are run at a refresh rate of either 120 or 150 hertz. Unfortunately Experiment 1 of the current study was run at 75 hertz due to an oversight. This led to a higher number of late display changes than is usual for this kind of study. However, we are confident that our data exclusion criteria ensured that the late changes did not contribute to the effects we report. Furthermore, this oversight was addressed in Experiment 2.

3.2.2.4. Materials and Design.

Forty-five pairs of words from which two new words could be created by both transposing and substituting the final and initial letter of the first and second words were embedded into sentences. The first word was always three-letters long, and the second either three- of four-letters long. The two target words were embedded beyond an invisible boundary, as word *n*+1 and word *n+2*, and had their parafoveal previews manipulated. Participants received either a correct preview (e.g. *hot pan*), a transposed letter preview (e.g. *hop tan*), or a substituted letter preview (e.g. *hob fan*). An example of the stimuli can be seen in Figure 3.1.

We faced a large level of constraint when preparing our stimuli. This was due to the fact that we needed to find pairs of words that fit into a sentential context, could be changed into a new pair of words via a single letter transposition, and changed into a further pair of words by substituting the same letters. Due to this, it was not always possible to closely match all three preview types on all possible characteristics for both word *n*+1 and word *n*+2. However, at a minimum it was necessary to ensure that no significant differences were present between the transposition and substitution conditions in terms of the frequency of the preview of either word *n*+1 or word *n*+2. In order to ensure that this was not the case we obtained frequency norms from the SUBTLEX corpus (Brysbaert & New, 2009) which is part of the English Lexicon Project (Balota et al., 2007). We conducted a paired t-test of the preview log frequency per million between the transposition (*n*+1 mean = 2.75, *n*+2 mean = 2.71) and substitution conditions (*n*+1 mean = 2.89, *n*+2 mean = 3.23). No significant differences were present for either the word *n*+1 preview (t (44) = -1.05, p = .30) or the word *n*+2 preview (t (44) = -1.38, p = .17). We also assessed whether our transposed letter previews differed from our substituted letter previews in terms of several other variables. There were no significant differences between the transposed letter and substituted letter previews in mean bigram frequency (word *n*+1 t (44) = 0.20, p = .84, word *n*+2 (t (44) = -0.48, p = .63), number of phonemes (word *n*+1 t (44) = 1.09, p = .28, word *n*+2 (t (44) = 1, p = .32) number or orthographic neighbours (word *n*+1 t (44) = 1.27, p = .21, word *n*+2 t (44) = -1.40, p = .17), neigbourhood frequency (word *n*+1 t (44) = 0.86, p = .39, word *n*+2 (t (44) = 0.18, p = .86), or number of phonological neigbours (word *n*+1 t (44) = -0.53, p = .60, word *n*+2 (t (44) = -0.39, p = .70).

3.2.2.5. Procedure.

Participants were presented with an information sheet and consent form upon arrival. A head rest was used to stabilize the reader in front of the eye-tracker. A three-point horizontal calibration grid was used, with an acceptance criterion of an average error below 0.33 degrees.

Each trial began with a central drift check followed by a gaze-contingent fixation cross in the position of the first character. The participant was re-calibrated if the cross did not trigger or the drift correct returned a value greater than 0.33 degrees. The participant was also re-calibrated at regular intervals. When the cross triggered a sentence appeared. Participants read for comprehension, and pressed a button once they had read each item. There were eight practice trials. The forty-five experimental items were mixed in with 64 filler items, 40 of which were part of another gaze-contingent study. On one third of the trials participants were shown a yes/no comprehension question, and answered using a game controller. Across all participants 94% of the questions were answered correctly. The experiment took approximately 30 minutes.

3.2.3.1. Results

Prior to analyzing our data we removed any fixations above 800ms, and merged any fixations below 80ms with any fixation less than 0.5 degrees of visual angle away, or any fixations below 40ms with any fixation within 1.25 degrees of visual angle. Trials in which the participant blinked while fixated on a critical region were excluded, as well as trials in which the display change executed early or more than 10ms after fixation onset on a post-boundary word (See Slattery, Angele, and Rayner, 2011). Altogether these exclusions accounted for 28% of the data. Finally, for each measure we removed any observations that were more than 3 standard deviations from the grand mean.

We examined several measures across several regions of interest in order to assess the effect of our manipulation. The measures we examined were first fixation duration (FFD; the duration of the first fixation on a word), gaze duration (GD; the sum of all fixations on a word from the first fixation until a saccade to another word), single fixation duration (SFD; the duration of a fixation when it is the only first pass fixation made on a word), go-past time (GP; the sum of all fixations from the first fixation in a region until a saccade was made to the right of the region), and skipping probability (SP; the probability that readers will skip a word in first pass reading). We examined these measures on word *n*, word *n*+1, and word *n*+2. In addition we also examined reading time measures on a composite region, consisting of both word *n*+1 and word *n*+2. Given that our preview manipulation extended across both of these words it does not seem unreasonable to assume that any effect of our manipulation may be more likely to appear in a complex manner across both of these words in a way that may not be detected through the analysis of the individual words. More generally, recent research has shown that preview manipulations on word *n*+2 may appear more clearly in composite, but not separate, regions especially when skipping rates of one of the regions is high (see Yu, Cutter, Yan, Bai, Fu, Drieghe, & Liversedge, 2016). The means and standard deviations of these measures are displayed in Table 3.1.

To analyze the data we constructed Linear Mixed Models (LMMs) using the lme4 package (Bates, Maechler, & Bolker, 2012) in R (2013). LMMs retain greater statistical power than ANOVAs in unbalanced designs (see Baayen, 2008) and so are well-suited to analyzing boundary studies, in which trials are often excluded due to technical errors. The preview condition was treated as a fixed factor. Helmert contrasts were used in order to compare 1) the identity preview condition to both the transposition and substitution preview conditions simultaneously and 2) to compare the transposition preview to the substitution preview. Subject and items were treated as random factors, with both random intercepts and slopes in accordance with Barr, Levy, Scheepers, and Tily (2013). In some cases the slopes were removed due to either a failure to converge or when the correlations in the random structure were equal to 1 or -1 which indicates overparametrization of the model. The beta values, standard errors, and t-values of the contrasts are displayed in Table 3.2.

3.2.3.2. Word *n*.

There were no significant effects of our manipulation on word *n*.

3.2.3.3. Word *n*+1.

There were significant identity preview effects across all fixation time measures on word *n*+1, such that participants would fixate this word for less time given an identity preview, relative to the two incorrect preview types. However, there were no significant differences between the transposition and substitution preview in any measure.

Furthermore, the direction of any numerical differences between these two conditions varied from measure to measure.

Table 3.1

Fixation Time Measures (in Milliseconds) and Fixation Probability for all Target Regions in Experiment 1. Standard Deviations are Presented in Parentheses.

Preview	Word *n*	Word *n*+1	Word *n*+2	Whole region
		First fixation duration		
Identity	220(64)	247(82)	237(84)	247(82)
Transposition	216(60)	261(87)	245(92)	264(91)
Substitution	214(60)	261(86)	249(92)	264(92)
		Gaze duration		
Identity	230(74)	259(93)	252(99)	383(183)
Transposition	227(72)	278(97)	276(123)	451(204)
Substitution	225(74)	280(98)	275(117)	446(200)
		Go-past time		
Identity	236(81)	272(104)	266(112)	411(189)
Transposition	235(79)	297(109)	301(139)	485(198)
Substitution	232(81)	293(107)	299(132)	486(205)
		Single fixation duration		
Identity	222(66)	251(84)	243(90)	256(90)
Transposition	214(55)	275(89)	241(93)	298(109)
Substitution	215(60)	277(90)	259(95)	309(100)
		Skipping probability		
Identity	.17(.38)	.30(.46)	.28(.45)	-
Transposition	.18(.39)	.30(.46)	.22(.42)	-
Substitution	.15(.36)	.30(.46)	.23(.42)	-

3.2.3.4. Word n+2.

There were significant identity preview effects on word *n*+2 in gaze duration, go-past time, and skipping probabilities, but not in first or single fixation durations. In the fixation time measures this effect was facilitative, such that participants fixated word *n*+2 for less time following an identity preview. However, in the skipping probabilities participants were more likely to skip word *n*+2 given an identity preview as opposed to either false preview. There was a significant difference between the transposition and substitution preview in single fixation durations, such that participants would fixate word *n*+2 for less time following a transposition relative to substitution preview.

Table 3.2.

Fixed Effects Estimates, Standard Errors, and t-values from the LMM contrasts for all Measures across All Regions in Experiment 1.

	First fixation duration			Gaze duration			Go-past time			Single fixation duration			Skipping probability		
Effect	*b*	SE	t	*b*	SE	t	*b*	SE	t	*b*	SE	t	*b*	SE	z
Word *n*															
Identity effect	0.01	0.02	0.72	0.01	0.02	0.59	0.01	0.02	0.28	0.02	0.02	1.02	0.02	0.15	0.12
Transposition effect	-0.01	0.02	-0.45	-0.01	0.02	-0.37	-0.01	0.02	-0.28	-0.01	0.02	0.43	-0.28	0.18	-1.59
Word *n*+1															
Identity effect	**-0.06**	**0.02**	**-2.90**	**-0.08**	**0.03**	**-3.27**	**-0.10**	**0.03**	**-3.51**	**-0.10**	**0.03**	**-3.27**	0.01	0.13	0.06
Transposition effect	0.00	0.02	0.10	0.01	0.02	0.24	-0.01	0.03	-0.55	0.02	0.03	0.62	-0.00	0.15	-0.03
Word *n*+2															
Identity effect	-0.05	0.03	-1.67	**-0.08**	**0.03**	**-3.21**	**-0.12**	**0.03**	**-4.11**	-0.03	0.03	-1.29	**0.31**	**0.13**	**2.47**
Transposition effect	0.02	0.03	0.73	0.01	0.03	0.28	0.01	0.03	0.31	**0.07**	**0.03**	**2.23**	0.04	0.15	0.25
Composite region															
Identity effect	**-0.07**	**0.02**	**-3.88**	**-0.17**	**0.03**	**-5.51**	**-0.19**	**0.03**	**-6.52**	**-0.18**	**0.04**	**-4.74**	-	-	-
Transposition	0.00	0.02	0.03	-0.01	0.03	-0.20	-0.00	0.03	-0.15	0.03	0.05	0.48	-	-	-

Significant terms are presented in bold.

3.2.3.5. Composite Region.

In our composite region we observed significant identity preview effects across all fixation time measures. There were no significant differences between the transposition and substitution previews, and the direction of any numerical differences between these conditions varied across measures.

3.2.4. Discussion

In Experiment 1 we presented readers with an identity preview of word *n*+1 and *n*+2, a preview in which the final letter of word *n*+1 and the initial letter of word *n*+2 had been transposed, or a preview in which the final and initial letter of word *n*+1 and *n*+2 had been substituted for alternative letters. We hypothesized that if word *n*+1 and *n*+2 were processed in parallel we may observe a preview benefit of the transposition condition relative to the substitution condition. Furthermore, we hypothesized that we would observe a preview benefit of the identity preview relative to the other two preview types, regardless of whether words are processed in serial or parallel. While we observed significant benefits of our identity preview condition relative to the two alternative preview types, we found very little evidence that word *n*+1 and *n*+2 were processed in parallel.

Across the three regions and five measures that we analyzed there was a significant difference between our transposition and substitution condition in just one measure, in one region. This effect was in single fixation durations, with word *n*+2 being fixated for 18 milliseconds less following a transposition preview as opposed to a substitution preview. It is worth discussing this effect in the context of the rest of our data in order to assess whether it was part of a larger, more meaningful trend, rather than simply being a spurious effect. We consider it to be the latter. Across all of our other measures and regions there was no consistent trend towards this effect, with there being many instances of the opposite trend being present. Had there been a genuine effect of the manipulation we would at least have expected some sort of consistent trend across measures, even if this generally failed to reach significance.[8]

[8] Subsequent exploratory data analysis of this transposition effect also indicated that it depended on the specific criterion used for data cleaning. For example, in a preliminary analysis in which outliers were removed on the basis of an observation being 2.5 standard deviations away from the mean for a participant per condition as opposed to the grand mean, this effect was only marginally significant. The

Given that our effect only appeared in single fixation durations it is worthwhile considering the proportion of trials that this measure represented. There are several ways to calculate this. The first way of looking at this issue is in terms of how representative this measure is of the processing of word *n*+2, and this can be considered in terms of how often participants would fixate this word only once, relative to more than once. In the case of our identity condition this occurred on 75% of trials when participants fixated word *n*+2 at all, whereas in both our transposition and substitution conditions participants would make a single fixation on word *n*+2 on 63% of the trials on which they fixated this word at all. An alternative way of calculating the proportion of single fixations on word *n*+2 is as a proportion of how often a participant fixated anywhere in our composite region. In this case, there was a 55%, 49%, and 50% probability of making a single fixation on word *n*+2 in the identity, transposition, and substitution conditions, respectively. While this is not a trivial proportion of trials, neither do I consider it to be more representative of the effect of our manipulation than all of the other measures which were calculated on the basis of every instance in which readers fixated either of the target words.

As well as being of the opinion that our other measures may be more representative of reading behaviour, there are statistical considerations that should be made when interpreting the effect in single fixation durations on word *n*+2. As outlined in section 3.2.3.1 we excluded data on the basis that participants either blinked, or there was an error in the execution of a display change, leading to the loss of 28% of our data. When taking this into consideration in the calculation of how much of the total data we have available in terms of assessing single fixation durations on word *n*+2, we are left with 36% of our data. While LMMs are generally robust against missing data values, it may well be the case that our data set was unbalanced to such an extent that artifacts emerged in this measure due to subject or item effects. In order to explore this possibility we attempted to assess the effect in single fixation duration on the basis of a dataset in which we did not exclude trials on the basis of the criteria discussed above. In this dataset, while there was a 7ms trend for the transposition effect, this did not come close to reaching significance in our LMM contrast (b = 0.04, SE = 0.03, t = 1.20). Furthermore, we then proceeded to assess this effect as we gradually excluded more and more data on the basis of the criteria listed earlier. We did not observe a significant difference between our transposition and preview condition when

dependence of this effect on precise data cleaning parameters, combined with the absence of this effect in first fixation duration, gaze duration, and go-past time, leads us to believe that it is most likely spurious in nature.

we excluded trials in which the display change triggered early (*b* = 0.05, SE = 0.04, *t* = 1.51), or when we also excluded blinks (*b* = 0.06, SE = 0.03, *t* = 1.60), or when we also excluded late display changes (*b* = 0.06, SE = 0.04, *t* = 1.71). Thus, when we were able to assess this effect using a greater proportion of our data, it was no longer significant. Furthermore, all of these models included random slopes, whereas the model reported in Table 3.2 did not, due to correlations of 1 in the random structure. Considering that part of the problem with omitting random slopes from an LMM is that this can lead to a high number of Type I errors (Barr, Levy, Scheepers, & Tily, 2013), it may well be that this effect would not have been significant in the reported model had we been able to include random slopes. None of this is to say that this effect should be completely ignored; rather, it may be wise to attempt to see if this effect is replicated under similar circumstances. Fortunately, Experiment 2 of the current chapter allows us to do this.

The results of Experiment 1 are in line with those of a recent investigation of Chinese reading (Gu & Li, 2015). In this study readers were presented with parafoveal previews of a four-character target region. These four characters would either form a single four-character word, or two two-character words. The parafoveal preview of this region would either be an identity preview, a preview in which the second and third character had been transposed, or a preview in which the second and third character were substituted; in the case of the two incorrect preview types the target region neither formed a single four-character word nor any two-character words. Depending upon whether the target region formed a single word or two words the transposition would either be within a word, or between two words. When the target region formed a single word the transposition preview led to significantly faster reading times than the substitution preview, similarly to the effects that occurred in Johnson’s (2007) study in English. When the target region formed two separate words, however, there were no differences between the transposition and substitution previews, a finding similar to the results in the current experiment. Our interest in the current study was primarily in whether orthographic information from one word in the parafovea can influence the processing of an adjacent parafoveal word. Given this, it is interesting to note that inter-word transposition effects do not seem to occur either in a language where parafoveal-on-foveal effects tend to be quite rare (i.e. English), or in Chinese where such effects are far more common (e.g. Yan, Kliegl, Shu, Pan, & Zhou, 2010; Yan, Richter, Shu, & Kliegl, 2009; Yan & Sommer, 2015). It is also interesting to note that such effects failed to occur regardless of whether word spacing was or was not present.

3.3.1. Experiment 2

Experiment 1 was primarily based upon prior research demonstrating transposed letter effects in parafoveal previews. However, there is a second phenomenon from isolated word recognition research which our preview manipulation can be viewed in relation to, referred to as letter migration errors (Davis & Bowers, 2004; Mozer, 1983; Treisman & Souther, 1986). This is an effect whereby presenting two or more words concurrently for a limited duration leads to the perception of an illusory word, which was a combination of the two presented words. For example, having seen *line* and *love* participants report the word *lone* as being present. This is referred to as a letter migration error since the misperception involves a letter from one word (e.g. the *o* in *love*) migrating to the other word to form the illusory word.

These effects have been reported across a number of tasks. Mozer (1983) briefly presented participants with a target word (e.g. *line*) paired with a context word (e.g. *love*). Once the words had disappeared a probe appeared where one of these words had been, and participants were required to name this word. When participants made an error it was more likely to involve reporting a conjunction of the two presented words (e.g. *lone, live*, or *love*) than a word that was not a conjunction (e.g. *lane, lice*, or *lace*). Davis and Bowers (2004) extended this finding by demonstrating that the effect can also occur across letter positions (e.g. *step* and *soap* resulting in the response *stop*). Treisman and Souther (1986) observed letter migration effects in a search task. They instructed participants to search an array of four words for a pre-specified target word, which could either be present or absent. The array was only briefly presented, in order to necessitate the distribution of attention across all words. On trials in which the target was absent, its constituent letters could either be distributed across the words in the array (e.g. *dab, dam, say, hay* for the target *day*) or were not present (e.g. *bud, bug, bun, bus* for the target *but*). Participants falsely responded that the target word was present more often in the former than the latter condition. This suggests that orthographic information extracted from multiple words in the array was fed into the lexicon simultaneously, in order to activate the lexical representation of the target word.

One interpretation of these findings is that they occur due to the distribution of attention across the multiple presented word forms. Presumably, the presented words must be orthographically encoded simultaneously due to the short exposure durations used in these studies. These multiple word forms may then proceed to feed activation into the lexicon in parallel, and as a result both sets of orthographic information activate lexical representations at the same time. Indeed, Davis and Bowers (2004) convincingly argued that the locus of these effects is most likely at the level of the lexicon. Consequently, a word which shares the orthography of all of the displayed words sometimes receives enough activation to lead to the false identification of an illusory word. In light of this, these letter migration errors may have interesting implications for the debate between serial and parallel approaches to lexical processing during reading. These effects suggest that when two words are lexically processed in parallel we should indeed expect orthographic information from one word to influence the processing of the other, in the way that Reichle et al. (2009) proposed.

Contrary to this suggestion, we observed no such effects in Experiment 1. This null effect could be due to one of at least two reasons. It could be that unlike in somewhat artificial, time-limited, isolated word recognition tasks, words are not typically processed in parallel during normal reading. However, there is an alternative possibility. In Experiment 1 our transposition was always made between the final letter of one word, and the initial letter of another. While Davis and Bowers (2004) showed that letter migration effects occur between letters from different positions within a word, this did not include the initial letter of a word. As such, we may have failed to observe an effect in Experiment 1 due to the within-word location of our manipulation, rather than our target words being lexically processed in serial order. We explored this possibility in Experiment 2 by making our transpositions between equivalent letter positions in the two words (e.g. *pit hop* rather than *hip top* as a transposition preview of *hit pop*). Our hypotheses here were very similar to those of Experiment 1, in that we predicted that if word *n*+1 and word *n*+2 are typically processed in parallel then we may observe a preview benefit in the transposition condition relative to the substitution condition.[9] On the other hand, if word *n*+1 and word *n*+2 are processed in serial order then we would only expect an identity preview benefit, with no difference between the transposition and substitution condition.

[9] While our manipulation in Experiment 2 could be more accurately described as a migration preview than a transposition preview we continue to use the latter terminology for the sake of consistency with Experiment 1.

3.3.2.1. Method

3.3.2.2. Participants.

Sixty-six native English speakers with normal or corrected to normal vision participated. An additional 6 participants were tested but removed from the analysis due to noticing more than five display changes. Participants were rewarded with research credits.

3.3.2.3. Apparatus.

The same apparatus was used as in Experiment 1, with the only change being that the monitor was running at a refresh rate of 120 hertz.

3.3.2.4. Materials and Design.

Thirty-three pairs of words from which two new words could be created by transposing or substituting two letters between the words were embedded in sentences. While the transposition could be made for any letter position in the word, it would always maintain the letters' within word positions. Both words were always three letters long. The two target words were embedded beyond an invisible boundary, as word *n*+1 and word *n+2*, and had their parafoveal preview manipulated. Participants received either a correct preview (e.g. *hit pop*), a transposed letter preview (e.g. *pit hop*), or a substituted letter preview (e.g. *fit cop*). Once again, we obtained frequency norms from the SUBTLEX corpus (Brysbaert & New, 2009) which is part of the English Lexicon Project (Balota et al., 2007). No significant differences were present between the transposition (*n*+1 mean = 3.26, *n*+2 mean = 3.14) and substitution condition (*n*+1 mean = 3.19, *n*+2 mean = 4.01) in terms of the log frequency per million of word *n*+1 (t (32) = -0.29, p = .77) or word *n*+2 (t (31) = -1.42, p = .16). Once again, there were no significant differences between the transposed letter and substituted letter previews in mean bigram frequency (word *n*+1 t (32) = 0.99, p = .33, word *n*+2 (t (31) = -0.91, p = .37), number of phonemes (word *n*+1 t (32) = 1.79, p = .08, word *n*+2 (t (31) = 1, p = .33), number or orthographic neighbours (word *n*+1 t (32) = 0, p = 1, word *n*+2 t (31) = -0.17, p = .86), neighbourhood frequency (word *n*+1 t (32) = 1.61, p = .12, word *n*+2 (t (31) = -1.35, p = .19) or number of phonological neighbours (word *n*+1 t (32) = 1.51, p = .14, word *n*+2 (t (31) = -0.55, p = .59).

3.3.2.5. Procedure.

For the most part the experimental procedure was identical to Experiment 1. The only changes were that the acceptance criterion for the calibration and drift check was lowered to 0.25 degrees and the gaze-contingent fixation cross was replaced with a second drift check. The thirty-three experimental items were mixed in with 43 fillers items. Across all participants 95% of the comprehension questions were answered correctly. The experiment took approximately 20 minutes.

3.3.3.1. Results

We used the same data exclusion criteria in Experiment 2 as in Experiment 1, accounting for a removal of 16% of the data gathered. We computed the same measures across the same regions of interest as in Experiment 1, and constructed LMMs to perform the same contrasts. Means and standard errors of the computed measures are presented in Table 3.3 and the output from our LMMs are shown in Table 3.4.

3.3.3.2. Word *n*.

No significant effects of our preview manipulation were observed during fixations on word *n*.

3.3.3.3. Word *n*+1.

Participants' first fixation durations, gaze durations, go-past times, and single fixation durations were significantly shorter when given an identity preview of word *n*+1 and word *n*+2 than a preview including either a letter transposition or substitution. However, we observed no significant differences between the transposition and substitution condition. Furthermore, any numerical differences were in the opposite direction to what was hypothesized.

Table 3.3

Fixation Time Measures (in Milliseconds) and Fixation Probability for all Target Regions in Experiment 2. Standard Deviations are Presented in Parentheses.

Preview	Word *n*	Word *n*+1	Word *n*+2	Whole region
		First fixation duration		
Identity	223 (59)	257 (89)	245 (83)	253 (87)
Transposition	219 (58)	274 (101)	260 (95)	272 (101)
Substitution	224 (65)	270 (92)	256 (91)	266 (94)
		Gaze duration		
Identity	257 (88)	272 (101)	279 (119)	426 (216)
Transposition	251 (92)	297 (116)	296 (129)	470 (224)
Substitution	253 (96)	291 (109)	292 (127)	472 (232)
		Go-past time		
Identity	269 (97)	285 (108)	297 (137)	464 (229)
Transposition	267 (102)	314 (119)	326 (147)	517 (229)
Substitution	267 (104)	304 (113)	327 (148)	520 (241)
		Single fixation duration		
Identity	229 (57)	262 (91)	250 (84)	263 (89)
Transposition	222 (51)	291 (97)	267 (95)	302 (91)
Substitution	225 (62)	280 (92)	271 (93)	302 (103)
		Skipping probability		
Identity	0.10 (0.30)	0.26 (0.44)	0.26 (0.44)	-
Transposition	0.09 (0.29)	0.28 (0.45)	0.24 (0.43)	-
Substitution	0.07 (0.26)	0.29 (0.45)	0.27 (0.44)	-

3.3.3.4. Word *n*+2.

There was a significant identity preview benefit during fixations on word *n*+2 in both go-past times and single fixation durations, as well as a marginal effect in first fixation durations. Once again, there were no significant differences between the transposition and substitution condition, and the direction of any numerical differences was inconsistent between measures.

3.3.3.5. Composite Region.

We observed significant identity preview benefit in first fixation durations, gaze durations, go-past times, and single fixation durations. However, no significant differences were observed between the transposition and substitution condition.

Table 3.4.

Fixed Effects Estimates, Standard Errors, and t-values from the LMM contrasts for all Measures across All Regions in Experiment 2.

	First fixation duration			Gaze duration			Go-past time			Single fixation duration			Skipping probability		
Effect	*b*	SE	t	*b*	SE	t	*b*	SE	t	*b*	SE	t	*b*	SE	z
Word *n*															
Identity effect	0.01	0.01	0.63	0.03	0.02	1.47	0.02	0.02	0.89	0.02	0.02	1.43	0.21	0.18	1.17
Transposition effect	0.01	0.02	0.75	0.00	0.03	0.06	-0.00	0.02	-0.14	0.00	0.02	0.08	-0.21	0.22	-0.95
Word *n*+1															
Identity effect	**-0.05**	**0.02**	**-2.51**	**-0.08**	**0.02**	**-3.41**	**-0.08**	**0.02**	**-3.52**	**-0.08**	**0.02**	**-3.46**	-0.14	0.12	-1.22
Transposition effect	-0.01	0.03	-0.32	-0.01	0.03	-0.36	-0.02	0.03	-0.82	-0.05	0.03	-1.46	0.02	0.14	0.89
Word *n*+2															
Identity effect	-0.04	0.02	-1.58	-0.04	0.03	-1.54	**-0.09**	**0.03**	**-3.20**	**-0.06**	**0.03**	**-2.22**	0.05	0.12	0.47
Transposition effect	-0.01	0.02	-0.54	-0.01	0.03	-0.38	0.00	0.03	0.02	0.01	0.03	0.36	0.15	0.14	1.11
Composite region															
Identity effect	**-0.05**	**0.02**	**-2.54**	**-0.10**	**0.03**	**-3.32**	**-0.13**	**0.03**	**-5.08**	**-0.16**	**0.04**	**-3.99**	-	-	-
Transposition	-0.01	0.02	-0.68	0.00	0.03	0.11	0.00	0.03	0.00	0.02	0.05	0.38	-	-	-

Significant terms are presented in bold.

3.3.4. Discussion

Experiment 2 was conducted in order to ensure that the null effect observed in Experiment 1 was due to parafoveal orthographic information extracted from word *n*+1 not influencing the lexical processing of word *n*+2 (and vice versa), rather than being a consequence of the transposition being made between the final and first letter of the two words. Our data suggest that it is indeed the case that orthographic information from one word in the parafovea does not influence the processing of another parafoveal word. Once again we observed significant identity preview benefits across the majority of measures, but nothing that suggested the transposition preview facilitated target word processing to a greater extent than the substitution preview. Furthermore, it is worth noting that the effect that we observed in single fixation durations on word *n*+2 in Experiment 1 did not replicate in this experiment, lending support to our assertion that this effect was spurious. Thus, it seems fair to conclude that, unlike in isolated word recognition tasks, letter migration effects do not occur during natural reading.

Before proceeding it is important to consider the possibility that while a letter migration effect did not appear in our overall data, there may have been a sub-set of trials in which such effects were present. Specifically, it could be the case that these effects are sensitive to the amount of time participants spent fixated on the pre-boundary word. As mentioned above, letter migration effects tend to be observed in isolated word recognition tasks as a result of short exposure durations to the words. Furthermore, the number of errors made tends to increase with shorter exposure durations. There is a large level of variation between studies observing these effects in the exposure durations used. For example, Treisman and Souther (1986) used an exposure durations of 150–250ms, Davis and Bowers (2004) an average of 96ms or 71ms, and Mozer (1983) an average of 348ms, with a range of 262–425ms. It is difficult to determine an exact range of exposure durations at which migration effects may be observed during reading, with some of the times in the studies listed above being quite long (e.g. 348ms) relative to the average fixation (i.e. 250ms) and others relatively short (e.g. 96ms). Indeed, this range could be seen to suggest that the amount of time spent fixating prior to the boundary should have had very little influence on letter migration effects. Nonetheless, we conducted further analyses to investigate this possibility. In order to test this hypothesis, we examined whether including the last fixation duration prior to crossing the invisible boundary significantly improved the fit of our LMMs. We tested this as both a main effect in the model, and an interactive term with our preview contrasts. While including this measure as a main effect did significantly improve the fit of our model for several measures (all measures on word n+1, first fixation duration on the composite region),[10] there was no measure for which allowing the last pre-boundary fixation duration to interact with the preview contrast led to a significant improvement in the fit of the model. For completeness, we note that the improvement for the model for single fixation durations on word *n*+2 was marginal (p=.059), but the crucial interaction between pre-target fixation and the contrast between substituted and transposed preview was not close to being significant, (*b*=-0.03, SE=0.11, t=-0.31).

Likewise, a re-analysis of Experiment 1 showed a similar pattern in that including pre-fixation duration in the LMMs did not improve the fit of any of the measures with the exception of single fixation duration on word *n*+2, where there was a marginal improvement (p=.06). We still examined the model, which did show a significant interaction between the final pre-target fixation and the contrast between the transposed and substituted preview (*b*=0.25, SE=0.11, t=2.17). To further investigate this interaction we performed a median split of our data on the basis of final pre-target fixation duration. Our effects appeared to be driven by the fact that in the case of the identity and substitution preview, the single fixation duration on word *n*+2 increased alongside pre-target fixation duration, increasing from 229ms to 253ms and from 247ms to 274ms respectively, while in the transposition condition single fixation durations remained constant, staying at 240ms for short and long final pre-target fixation durations. Thus, not only was the improvement in our model merely marginal, it was driven by a trend in which the identity preview was less facilitative than a transposition preview with increased exposure durations. Between the improvement in the fit our model being marginal and the pattern of effects making little sense, it seems fair to consider them as not representing a meaningful finding.

One criticism we foresee of Experiment 2 that we wish to address before moving on is that word *n*+1 and word *n*+2 were not orthographically similar. This could be considered a problem, since prior work on letter migration effects has found that the effects tend to be observed when the presented words are similar (e.g. *step* and *soap* leading to the perception of *stop*; *dab, dam, say* and *hay* leading to the perception of *day*) but not when

[10] It could be argued that we should report these additive models in place of our basic models, due to the improved fit. However, the only reason we constructed these models was to use them as a baseline to compare our interactive model to, rather than there being a theoretically interesting reason to examine a main effect of pre-target fixation duration. As such, we see little point in reporting these models in addition to our main analyses.

the presented words are dissimilar (e.g. *step* and *frog* do not lead to the perception of *stop*). However, consideration of the theoretical explanation for this effect suggests that this should not have been an issue in our experiment. In typical studies of letter migration effects the participant's task is to name a single target word. As such, observing a letter migration effect in these studies not only depends upon the orthographic information from the two presented words causing activation in the lexicon in parallel, but also that the illusory word becomes lexically activated enough to lead to the identification of that word, rather than either of the two presented words. In the *step/soap* example above where the illusory word shares 75% of orthographic information with both presented words this level of activation is likely to be reached; three letters from each presented word would activate the illusory word's lexical representation. In contrast, in the *step/frog* example above, the two presented words are more dissimilar, and the illusory word shares 75% of orthographic information with one of the presented words, but only 25% of such information with the other. Under these circumstances it is unlikely that the illusory word would become sufficiently activated to be a serious candidate. Consequently, in a winner-takes-all task such as naming this word is unlikely to be identified. However, we do not consider this to be an issue in a natural reading study implemented using the boundary paradigm. Unlike in studies using a naming paradigm, preview benefit effects are not dependent upon the identification of a single lexical item so much as the activation of multiple lexical items which overlap with the previewed text on one of several abstract characteristics (i.e. orthography, phonology, or semantics). Any lexical item that is activated on the basis of this parafoveal information will be identified faster upon fixation. As such, in the current study any cases in which orthographic information from both parafoveal words was fed into the lexicon in parallel should have been enough to lead to a benefit of the migration preview relative to the substitution preview. However, in a naming task the activation of the target words from the migration preview would not need to exceed the activation of the words in the parafovea.

3.4. General Discussion

In the current study we investigated a prediction which should arguably be made by parallel accounts of lexical processing (Reichle et al., 2009). We investigated this by examining whether orthographic information from one parafoveal word influences the processing of another parafoveal word. We presented readers with parafoveal previews of words *n*+1 and *n*+2 in which a letter of each word had either been transposed to the other word, or been substituted with an alternative letter. We hypothesized that if orthographic information from both words was fed into the lexicon at the same time then we would observe significantly shorter fixation times given a transposition—rather than substitution— preview. However, while we observed significant identity preview benefits there was no notable effect of our transposition preview relative to our substitution preview. This was the case both when the preview manipulation affected the final character of word *n*+1 and the first character of word *n*+2 (Experiment 1) and when our manipulation affected the letter at the same letter position in both word *n*+1 and word *n*+2 (Experiment 2).

Our findings offer no support for the notion that word *n*+1 and word *n*+2 are processed simultaneously while still in the parafovea. As discussed earlier, if both word *n*+1 and word *n*+2 feed orthographic information into the lexicon in parallel then we may expect to see orthographic information from these two words influencing the processing of each other. Had we observed such an effect, it would have been highly compelling evidence that multiple words within the perceptual span are processed in parallel during natural reading, and thus for a model such as SWIFT (Engbert et al., 2005; Schad & Engbert, 2012). It should be noted, however, that this model does not explicitly define how the parallel lexical processing of multiple words should occur. As such it is difficult to be sure as to whether this model would even predict an effect of our manipulation.

Attempts have recently been made to design a model of multiple-word reading, albeit outside of the context of eye movement control (Grainger, Mathôt, & Vitu, 2014; see also Grainger, Dufau, & Ziegler, 2016). While our findings do not necessarily pose a problem for the general concept of parallel lexical processing, they do have implications for this model. In their model Grainger et al. use an open-bigram coding scheme, whereby both contiguous (e.g. *ro* in *rock*) and non-contiguous (e.g. *rc* in *rock)* pairs of letters within a word activate bigrams, which in turn feed activation up to whole word representations. Crucially, in addition to the letters within a word being able to form bigrams with each other, they are also able to form bigrams with the space preceding and following the word. While letters are not able to form bigrams with letters from spatially adjacent words, the bigrams that are activated from several words within the perceptual span all feed activation up to a single whole word representation, and it is through this mechanism that orthographic information from one word is able to influence the processing of another word. As such, if the bigrams produced by one word overlap with the bigrams of a spatially adjacent word, then word identification should be facilitated relative to when the words do not share bigrams. Our results contradict this position. In both of our experiments the letters available in the transposition preview should have led to the parafoveal activation of a greater number of bigrams from the target words than the letters available in the substitution preview. Take, for example, the case of the transposition preview *hop tan* for the target words *hot pan.* Here, the *t* in *tan* and the space at the beginning and end of this word would have activated the bigrams *#-t* (in which # represents a space) and *t-#*, which are part of the target word *hot*. Furthermore, the *p* in *hop* and the spaces at the beginning and end of this word would have activated the bigrams *#-p* and *p-#* from *pan*. Upon these bigrams feeding activation to the whole word representations they would increase the activation of our target words, due to no longer being specifically linked to a particular word in the perceptual span at this level of the model. In contrast, the *b* and *f* in the substitution preview *hob fan* would not activate any of these four bigrams by joining with the spaces on either side of each word. A difference of two bigrams per word does not seem trivial for three letter target words, which, in the coding system used by Grainger et al., would only contain nine bigrams in total. As such, our results pose a challenge for this approach, and may need to be taken into consideration in any future iteration of this model. It should be noted that while this model does not explicitly state that information is extracted from word *n*+2, it presumably should if attempting to account for both word *n*+1 and word *n*+2 orthographic preview effects using a single mechanism.

Although we do not consider our results to necessarily be evidence against the general idea of parallel processing, they do raise a number of interesting questions. Specifically, our findings have relevance for phenomena that are sometimes interpreted as evidence of parallel lexical processing. As discussed in depth earlier, it could be assumed that orthographic parafoveal-on-foveal effects are due to orthographic information from the parafoveal word either facilitating (e.g. Angele et al., 2013) or inhibiting (e.g. Inhoff et al., 2000) the processing of a spatially adjacent word. However, our findings call this explanation into question, with very little evidence of orthographic information from one word influencing the processing of an adjacent word. It could be argued that the main difference here is that in our study orthographic information from word *n*+2 would have had to influence the processing of word *n*+1 (and vice versa), as opposed to orthographic information from word *n*+1 affecting the processing of word *n*, as in prior studies. However, for this to be a plausible explanation for the difference, a parallel processing account would have to make a clear distinction between the way in which word *n* and word *n*+1 are processed in parallel relative to the way in which word *n*+1 and word *n*+2 are processed in parallel. No such distinction is currently made. It could be that the decrease in visual acuity further into the parafovea could go some way to explaining our lack of effect from the perspective of a parallel model, with information from word *n*+2 not being extracted quickly enough to influence the processing of word *n*+1; however, presumably by this same logic the model would predict no *n*+2 preview effects. Once again, it is difficult to see exactly how a parallel model could approach lexical processing in such a way as to account for existing patterns of effects in a plausible manner. We consider our findings to be more in line with the explanation of orthographic parafoveal-on-foveal effects that was proposed by Reichle, Pollatsek, and Rayner (2006), whereby unusual visual configurations pop-out from the page, and interfere with typical processing. In studies showing an effect of word *n*+1's orthographic characteristics on the processing of word *n* such unusual visual configurations are typically present, whereas in the current study there was nothing visually unusual about our manipulation, hence leading to the null effect.

In summary, the current study aimed to assess whether word *n*+1 and word *n*+2 are lexically processed in parallel during reading. This was tested by examining whether orthographic information from two words in the parafoveal preview affected the processing of each other. Our findings suggested this was not the case, with participants not gaining a benefit from a preview in which the correct orthographic information was present in the parafovea as part of the wrong words, relative to a preview in which the correct orthographic information was completely absent. While this result is not necessarily evidence against parallel lexical processing, it certainly needs to be taken into account in any future attempts to model the parallel processing of multiple words and in the interpretation of phenomena such as parafoveal-on-foveal effects.

Chapter 4

Reading Sentences of Uniform Word Length: Evidence for the Adaptation of the Preferred Saccade Length during Reading

At the time of thesis submission the current Chapter has been revised on the basis of reviews from an initial submission for publication. This revision will soon be re-submitted.

4.1.1. Introduction

During reading saccadic eye movements are made in order to fixate words in high acuity foveal vision (see Rayner, 1998; 2009 for reviews). Typically these saccades will be targeted towards and originate from word objects of variable length. The length of the words within a sentence influences eye movement control during reading in several ways. First, the eyes will tend to land further into a long word than into a short word, necessitating saccades of differing lengths to move further when targeting long relative to short words (McConkie, Kerr, Reddix, & Zola, 1988). Secondly, the probability of skipping a word is dramatically affected by word length, with increased skipping of short than long words (see Brysbaert, Drieghe, & Vitu, 2005). In the current paper we investigated the effect of removing within sentence word length variability on all of these components of eye movement control. While word length also affects reading times, the focus of the current paper is upon spatial, rather than temporal, aspects of eye movement control.

4.1.2. Saccadic Targeting and the Systematic Range Error

It is generally agreed that progressive, inter-word saccades are targeted towards the centre of an upcoming word in spaced alphabetic languages. This location is considered the *optimal viewing position* for word recognition, with results from isolated word recognition studies showing that response latencies increase the further a fixation is from a word's centre (O'Regan & Jacobs, 1992; O'Regan, Lévy-Schoen, Pynte & Brugaillère, 1984). Furthermore, in natural reading the probability of refixating a word increases with the distance of the first fixation from the central character (McConkie, Kerr, Reddix, Zola, &

Jacobs, 1989; Rayner, Sereno, & Raney, 1996). Despite this, fixation landing positions tend to be normally distributed across the whole word (Rayner, 1979), indicating that the programming and execution of eye movements are not error-free, resulting in the eyes not always landing on the intended landing position. The peak of this normal distribution is slightly to the left of the word centre, and is referred to as the *preferred viewing location* (Rayner, 1979). Thus, while the eyes are targeted towards a word's centre in English reading, they tend to fall a small amount short of this, landing slightly to the left of the centre.

McConkie, Kerr, Reddix, and Zola (1988) investigated factors influencing the distribution of initial fixation landing positions within a word. They showed that both the mean initial fixation location and the standard deviation of fixations around this position was partly determined by the distance of the prior fixation from the centre of the target word. Saccades launched from a greater distance resulted in a wider distribution of landing positions. McConkie et al. proposed that this occurred as a result of random motor error, with longer saccades being prone to increased motor error. More importantly for the current investigation, McConkie et al. also proposed a systematic saccadic range error during reading. They observed that English readers tended to overshoot the centre of a word for saccades launched from less than seven characters away from the centre, and undershoot for saccades launched from more than seven characters away. Thus, in English reading there is systematicity in the saccadic targeting system such that saccades tend to be seven characters in length, which we will refer to as the *preferred saccade length*. This preferred saccade length can be defined as the intended saccade length that is not biased to either under or overshoot its intended target. McConkie et al. found that for each additional character that the saccade launch site was from seven characters away from the target word centre, the mean landing position of the eye would shift approximately half a character before or beyond the centre of the word (see Nuthmann, Engbert, & Kliegl, 2005, Paterson, Almabruk, McGowan, White & Jordan, 2015, and Yan et al., 2014, for similar findings in German, Arabic, and Uighur respectively).

McConkie et al. derived the idea of a systematic saccadic range error with a preferred saccade length from Kapoula (1985; see also Kapoula and Robinson, 1986). Kapoula (1985) had participants make saccades towards a target that appeared between 2.7 and 9.5 degrees of visual angle from a central launch site. When the target was at the lower end of this range participants' saccades typically overshot the target, and in the upper end of the range they often undershot the target. In a follow-up investigation, half of the original participants completed a similar task, but with targets ranging from 7 to 21.9 degrees away. Once again, participants showed a clear range bias. Crucially, a target that was far and was thus undershot in the first task (i.e. targets 7-11 degrees away) became a relatively near target in the second, and was overshot. Kapoula's findings suggested that in low-level oculomotor tasks the range bias is rapidly adaptable, with the range of targets in a stimulus set determining the preferred saccade length. One aim of the current study was to determine whether the preferred saccade length is also adaptable during reading.

The systematic saccadic range error is implemented as an important parameter in the two most dominant models of eye movement control, the E-Z Reader model (Reichle, Rayner, & Pollatsek, 1999; Reichle, Rayner, & Pollatsek, 2003) and the SWIFT model (Engbert, Nuthmann, Richter, & Kliegl, 2005; Schad & Engbert, 2012), where the saccadic range error is vital for explaining landing positions within words. In both models the range error is programmed so that approximately 0.4 characters of under- or overshoot occurs for every character that a saccade was intended to move beyond or before the preferred saccade length. In E-Z Reader this preferred saccade length is seven characters, while in SWIFT it is 5.4 characters for progressive inter-word saccades. While the two models differ on the preferred saccade length, this is currently treated as a fixed parameter in both. The models are agnostic about how this preference may have developed, but it seems reasonable to suggest that they could both, in principle, be modified to allow for a flexible preferred saccade length without violating any of their key theoretical assumptions.

4.1.3. Word Length and Launch Site Effects on Word Skipping

Word length also has significant effects on word skipping. Vitu, O' Regan, Inhoff, and Topolski (1995) found that readers skip three, four, and five letter words approximately 60%, 45%, and 30% of the time, respectively. This effect is driven by the spatial extent of the word, rather than the number of letters. Hautala et al. (2011) embedded a four or six letter word within a sentence either presented in a font in which the six letter word subtended a greater visual angle than the four letter word, or a font in which both words subtended the same amount of visual angle. When the six letter words subtended a greater visual angle they were skipped less often than the four letter words. In contrast, when both word lengths subtended the same visual angle there was no significant effect of the number of letters on skipping (see also McDonald, 2006). Thus, the decision of whether or not to skip a word seems to be largely driven by a word's spatial extent. Importantly, words with a greater spatial extent will extend further into the parafovea, and thus will fall in lower acuity vision and will be less likely to be processed as fully as when they are directly fixated.

It is also worth considering whether the decision to skip a word may be partially determined by its relative length within a sentence, rather than simply its absolute length. It may be that the decision about which words to skip will partially be based upon whether certain words are likely to be of high informational value. One cue for informational value is word length. In normal sentences shorter words typically tend to be of lower informational value than longer words (Piantadosi, Tily, & Gibson, 2011), partly because many short words act as function words, though this also holds for content words. For instance when a word is of low informational value within a sentence an abbreviated form (e.g. *chimp* for *chimpanzee*) is more likely to be used (Mahowald, Fedorenko, Piantadosi, & Gibson, 2013). Furthermore, Mahowald et al. demonstrated that this relationship forms a part of people's abstract linguistic knowledge. It might, therefore, make sense for readers to skip shorter words, since these will less likely be of high informational value. In Hautala et al.'s (2011) study these cues were absent with six-letter words extending across the same spatial distance as four-letter words. The lack of a difference in skipping of short and long words in this study may have partially been due to the absence of relative length information. We will return to this issue below.

Word skipping is also affected by the systematic saccadic range error. A consequence of saccadic overshoot or undershoot is that a word will sometimes receive a mislocated fixation, such that a word that was not the intended saccade target is fixated (Nuthmann, Engbert, & Kliegl, 2005). For example, if a saccade is targeted towards a three letter word from the final character of the prior word, then saccadic overshoot may cause the eye to land on the following word. This would constitute an accidental skip. Through simulations, Nuthmann et al. were able to estimate the probability for words of a certain length either receiving a mislocated fixation or being accidentally skipped. While words of all lengths were approximately equally likely to be accidentally fixated or undershot, some were more likely than others to be accidentally skipped (see Table 6 on page 2212 of Nuthmann et al.). While the probability of accidentally skipping a three-letter word was 0.28 this decreased to 0.16 and 0.08 for four- and five-letter words. Thus, due to the saccadic range error short words will accidentally be skipped more often than longer words. Consequently, we may generate the prediction that if the preferred saccade length adapts to novel conditions, then we may also expect to observe a corresponding change in word skipping rates.

4.1.4. The Current Study

In the current study participants read four types of sentences. In three of these all of the words were of uniform length, with all words being three-letters long (referred to as U3; e.g. *The sad boy had not had any fun all day*), four-letters long (U4; e.g. *They went over some very hard sums last week*) or five-letters long (U5; e.g. *David often plays awful death metal music about Satan*). In a non-uniform condition, participants read sentences comprised of words with a mixture of these word lengths (NU; e.g. *Tim can often leave work about one hour early*). On average the words within these sentences were 3.94 letters long. Participants read each type of sentence in separate blocks, such that a participant would read thirty uniform sentences of three letter words, followed by thirty uniform sentences of four letter words, then thirty uniform sentences of five letter words, and thirty non-uniform sentences. The order of block presentation was counterbalanced between participants. We hypothesized that if the preferred saccade length which determines the systematic saccadic range error during reading is adaptable then it should vary between these different types of sentences. Across the different types of sentence the mean distance between saccade targets will systematically vary, with inter-word saccades of a mean of four, five, and six characters being required to move between the centres of words in uniform sentences consisting of three, four, and five letter words, respectively (i.e. the length of each word, in addition to a space). In the non-uniform sentences the distance between saccade targets will be more variable, but should on average be about five characters (i.e. 3.94 letters per word, in addition to a space). As such, we predicted that if the preferred saccade length was flexible and readers adjusted it contingent on the length of words in the text being read then it would be smallest when participants read sentences consisting of three letter words, slightly larger when reading the non-uniform sentences and the uniform sentences of four letter words, and larger still when reading uniform sentences of five letter words.

We also predicted significant effects of our manipulations on skipping rates. First, we expected standard word length effects, such that for sentences comprised of short words we would see increased word skipping. We also made a hypothesis contingent upon our saccadic adaptation effect. As mentioned, short words are accidentally skipped due to the systematic saccadic range error more often than long words. Presumably, if readers become more adept at making short saccades in the uniform sentences of three letter words then there should be less accidental skipping of three letter words in these sentences than in the non-uniform sentences due to saccadic overshoot. Experimentally determining whether a word was skipped accidentally or on purpose is not possible, and as such we can only assess this hypothesis on the basis of whether the preferred saccade length adapts, and whether this is accompanied by systematic variation in word skipping rates.

We were also curious to see if sentence uniformity had a further effect on word skipping due to the absence of relative word length cues within sentences. First of all, we expected to observe a substantial decrease in the skipping of three letter words appearing in a uniform compared to non-uniform sentence, simply due to the fact that maintaining a global skipping rate of 60% (Vitu et al,. 1995) across an entire sentence would involve encoding visual information from very few positions in that sentence. As such it seems unlikely that participants would skip three letter words as often in a sentence consisting exclusively of words that length, than when these words appeared in sentences of mixed word length. Furthermore, if readers do use relative word length information as a cue to the informational content of a word, and therefore the necessity for direct fixation, we may expect that while a three letter word may be more likely to be of low informational value in a non-uniform sentence, this would be less the case in a sentence comprised entirely of three letter words. Consequently, three-letter words may be less frequently skipped in uniform than non-uniform sentences. The converse may be true for five-letter words. In the case of non-uniform sentences the five letter words may be judged to potentially be of greater informational value than either three or four letter words. Consequently, they would be more likely to attract a direct fixation. In contrast, presumably in a sentence consisting of exclusively five letter words it would be likely that more of these five letter words would be function words with low informational value, and as such length may no longer be such a good cue for each five letter word being of high informational value. Consequently, less of these words may be directly fixated out of an assumption that they will be of high informational value.[11] In short, we may observe decreased skipping of three letter words and increased skipping of five letter words in uniform relative to non-uniform sentences.

[11] We feel it is important to point out that we realize that the probability that a word will be a content word (as opposed to a function word) is fixed and does not change in relation to word length context. Instead, we are suggesting that a uniform word length constraint changes the likelihood that function or content words of that length will be present in those sentences.

4.2.1. Method

4.2.2. Participants

Twenty-four students at the University of Southampton with normal or corrected to normal vision participated in return for course credits.

4.2.3. Apparatus

Participants' eye movements were monitored using an SR Research Eyelink 1000 system with a sample rate of 1000 hertz. Only the right eye was tracked. Sentences were displayed in black on a grey background, on a single line of a ViewSonic p227f CRT monitor. Viewing distance was 78cm, with 1° of visual angle being occupied by 2.9 characters of monospaced Courier font.

4.2.4. Materials and Design

Four sets of thirty sentences were created for the current experiment. In three of these sets all words in a sentence were of uniform length, such that the sentences were comprised entirely of three letter words, entirely of four letter words, or entirely of five letter words. The fourth set consisted of sentences made of a combination of three, four, and five letter words. A within subjects design was used, such that all participants saw all 120 sentences. The sentences consisted of between eight and twelve words.

Eighteen participants rated our sentences for naturalness on a scale from one (very unnatural) to five (perfectly natural). The mean rating for the uniform sentences of three, four, and five letter words, and the non-uniform sentences were 3.69, 3.45, 3.47, and 3.87, respectively. A one-way ANOVA revealed significant differences between the four conditions, $F(3, 51) = 23.46$, $p <.001$. Follow up t-tests comparing the non-uniform sentences to each uniform sentence type revealed that the non-uniform sentences were rated as being significantly more natural than all three uniform sentence types (NU vs. U3 $t(17) = -2.58$, $p < .01$; NU vs. U4 $t(17) = -5.78$, $p < .001$; NU vs. U5 $t(17) = -5.34$, $p < .001$). While these differences were significant it should be noted that they are fairly small.

We also examined the frequencies of the words making up our sentences, as a function of word length and uniformity. Frequencies were based on the Zipf scale introduced by van Heuven, Mandera, Keuleers, and Brysbaert (2014), and obtained from the SUBTLEX corpus (Brysbaert & New, 2009) within the English Lexicon Project (Balota et al., 2007). We conducted t-tests to examine whether the frequency of the words of one length in the uniform sentences was the same as words of that length in the non-uniform sentences. There were no differences in the frequency of five letter words appearing in a uniform (m = 5.02) relative to non-uniform (m = 5.09) sentence (t(109.5) = -0.62, p > .1), or of four letter words in a uniform(m = 5.60) relative to non-uniform (m = 5.75) sentence (t(185.4) = -1.49, p > .1). However, three letter words appearing in uniform sentences (m = 5.87) were significantly less frequent than three letter words appearing in a non-uniform (m = 6.38) sentence (t(112.9) = -3.65, p < .001).

Due to the emergence of these confounds we do not report reading time measures, which are highly likely to have been affected by these issues, although any interested readers can find both descriptive and inferential statistics on these measures in Appendix A. Furthermore, we make an effort to assess any potential effect of frequency in our analysis of skipping rates and saccadic targeting.

4.2.5. Procedure

Participants completed the experiment in two sessions across two days, each lasting twenty minutes. In each session participants were presented with two blocks of thirty experimental sentences, preceded by six practice trials. All sentences from a single condition were presented in the same block. We adopted a blocked design to maximize our chances of obtaining adaptation effects. The order of block presentation was counterbalanced between participants.

Upon arrival at the first session participants were presented with a consent form and information sheet. They were seated in front of the eye-tracker and a head rest was used for stabilization. A three point horizontal calibration grid was used, with an acceptance criterion of an average error below 0.25 degrees.

After a successful calibration the experiment began. Each trial began with two drift checks. The first drift check was in the centre of the screen, while the second was on the left in the position of the central character of the first word of the sentence. If either indicated more than 0.3 degrees of error the participant was recalibrated. After the drift checks had been completed the experimental sentence appeared. Participants were instructed to read for comprehension, and press a button once they had read the sentence. On 33% of the trials the experimenter read the participant a yes/no comprehension question, and the participant responded using a button box. Across all participants 91% of comprehension questions were answered correctly. At the end of the second session the experimenter asked the participant if they had consciously noticed the experimental word length manipulation. Only one participant had.

4.3.1. Results

We computed global and local measures pertaining to each theoretical issue outlined above. The global measures were calculated across whole sentences (see Table 4.1). The four sentence types were compared using linear mixed effects models, constructed using the lme4 package (Bates, Maechler, Bolker, & Walker, 2014) in R (R Development Core Team, 2013). Uniform sentences of four letter words were treated as the baseline, due to being the most appropriate condition to compare our non-uniform sentences with due to the similar mean word length, and to allow us to compare these sentences to uniform sentences with shorter and longer words. Model output is shown in Table 4.2. For all models we initially adopted a full random structure, treating both participants and sentences as random factors, with random intercepts and slopes (see Barr, Levy, Scheepers, & Tily, 2013). However, due to the models for some measures failing to converge, random slopes were sometimes removed.

Our local measures examined the effect of our manipulation on each individual word within our sentences. The interest area used for each word included the space preceding the word.[12] After removing data for the first or last word of a sentence, and cases of tracker loss, 21592 words were available for analysis (out of a full data set of 28176). Of these 21592 words, 17080 received a direct first pass fixation, allowing them to be used in the landing position analysis, and reading time analysis presented in Appendix B. For each measure that we examined we excluded values that were more than 2.5 standard deviations from the mean of each participant in each condition as outliers. The amount of data

---

[12] Some researchers may prefer to exclude fixations made on the space preceding a word when conducting a landing position analysis, since these fixations did not technically land on the word. While our approach is in line with the classic papers on landing positions, we also conducted all analyses involving landing position as a dependent or predictor variable on a restricted dataset in order to assuage any concerns arising from this point. In this restricted dataset we excluded words in which the first fixation was on the space preceding the word. The differences between the two analyses were minimal. Excluding trials in which the initial fixation was on the space preceding a word also allowed us to scale landing position relative to word length. This transformation also made minimal difference to the outcome of our analyses. We also acknowledge that typically the term 'target word' would be used to refer to specific words that had been manipulated within a sentence for certain characteristics. However, for the purpose of the current investigation we use this tem to refer to any individual word within our sentences.

excluded due to this final criterion varied between measures, with a maximum of 3.2% of the remaining data being lost. All models described treated both subjects and individual words as random factors, with both random slopes and intercepts. Similar to our global analyses, the models for some measures failed to converge with the full random structure, and as such random slopes were sometimes removed.

4.3.2.1. Saccadic Targeting

At a global level our manipulation had a clear effect on saccade metrics (see Table 4.1). The mean saccade amplitude increased with the mean word length in a condition, with the shortest mean saccade length being observed in the uniform sentences of three letter words, and the longest in the uniform sentences of five letter words. Unexpectedly, participants made longer saccades in the non-uniform sentences compared to the uniform sentences of four letter words. This is most likely due to differences between these conditions in word skipping, which will be discussed in detail below. The linear mixed-effects models (see Table 4.2) revealed that the differences between the uniform sentences of four letter words and the other conditions were significant. Clearly as the distance between the centres of two words decreased, so did the amplitude of saccades moving between them.

Table 4.1
Mean Global Reading Measures for each Sentence Type. Standard Deviations are Presented in Parentheses.

	U3	U4	U5	NU
MSA (degrees)	1.68(0.40)	1.88(0.40)	2.13(0.50)	2.00(0.50)
MSA (letters)	4.87(1.16)	5.45(1.16)	6.18(1.45)	5.80(1.45)
PWF	0.77(0.12)	0.82(0.12)	0.89(0.10)	0.78(0.12)

Note. MSA= mean saccade amplitude; PWF= the proportion of words fixated in a sentence. U3, U4, and U5 refer to uniform sentences of three, four, and five letter words respectively, and NU to non-uniform sentences.

Given that saccade amplitudes varied between conditions, we next considered whether this affected the systematic range error. Primarily, we were interested in seeing whether the preferred saccade length was constant across our different types of uniform sentences, or whether it increased systematically with the length of words within these sentences. Similarly, as a control, we were interested in seeing whether the preferred saccade length for words of different lengths within non-uniform sentences was near constant, as in McConkie et al.'s (1988) original investigation. If the preferred saccade length did adapt then it should have increased alongside the length of words between the three different uniform sentence types, while any differences between the word lengths in non-uniform sentences should be minimal. Given this theoretical question, and that the preferred saccade length is the distance between the saccade launch site and the centre of the word on which the eyes land at which saccades are most accurate, all the analyses below calculated the saccade launch site relative to the centre of the word to which a saccade was made.[13]

Table 4.2

Linear Mixed Effects Models for Global Reading Measures.

	MSA			PWF		
Model	*b*	*SE*	*T*	*b*	*SE*	*t*
Intercept	0.61	0.03	18.28	0.82	0.02	47.84
U3	**-0.12**	**0.02**	**-5.92**	**-0.04**	**0.01**	**-5.33**
U5	**0.12**	**0.02**	**6.82**	**0.07**	**0.01**	**6.69**
NU	**0.06**	**0.02**	**2.39**	**-0.04**	**0.01**	**-3.41**

Note. Significant factors are presented in bold. MSA= mean saccade amplitude in degrees; PWF= the proportion of words fixated in a sentence. U3, U4, and U5 refer to uniform sentences of three, four, and five letter words respectively, and NU to non-uniform sentences.

4.3.2.2. Linear Mixed Effects Models. As a first step in our analysis, we constructed linear mixed effects models with the fixation landing position within each word as a dependent variable. The landing position was calculated relative to the word centre, such that a landing position of 0 corresponded to a saccade landing perfectly in the centre of the word, while landing positions of 1 or -1 corresponded to a saccade over- or undershooting the centre of the word by one character, respectively. This dependent variable was examined using a linear mixed-effect model with word length, sentence uniformity, and saccade launch site as fixed factors, in addition to two- and three-way interactions between these factors. Words with a launch site greater than 12 characters were not included in this analysis due to visual inspection of the launch site distributions identifying them as outliers, and the fact that it is unlikely that the fixated word was in high

[13] Many papers investigating the saccadic range error also include an analysis in which the launch site is calculated relative to the space before the target word. We did also construct a linear mixed effect model with this measure calculated from the start of the word, and calculated the preferred saccade length values from this model. Unsurprisingly, the same pattern of effects emerged.

enough acuity vision for readers to make an accurate estimate of the word centre. This exclusion accounted for 1.63% of the remaining data.

Table 4.3

Linear Mixed Model Analyses for Fixation Landing Position Data.

	LP		
Model	*b*	*SE*	*t*
Intercept	0.01	0.23	0.05
SU	0.22	0.36	0.59
WL	**0.64**	**0.05**	**12.51**
SU*WL	-0.09	0.09	-1.04
LS	**-0.19**	**0.04**	**-5.11**
LS*WL	**-0.07**	**0.01**	**-10.81**
LS*SU	**0.17**	**0.05**	**3.15**
LS*SU *WL	**-0.03**	**0.01**	**-2.48**

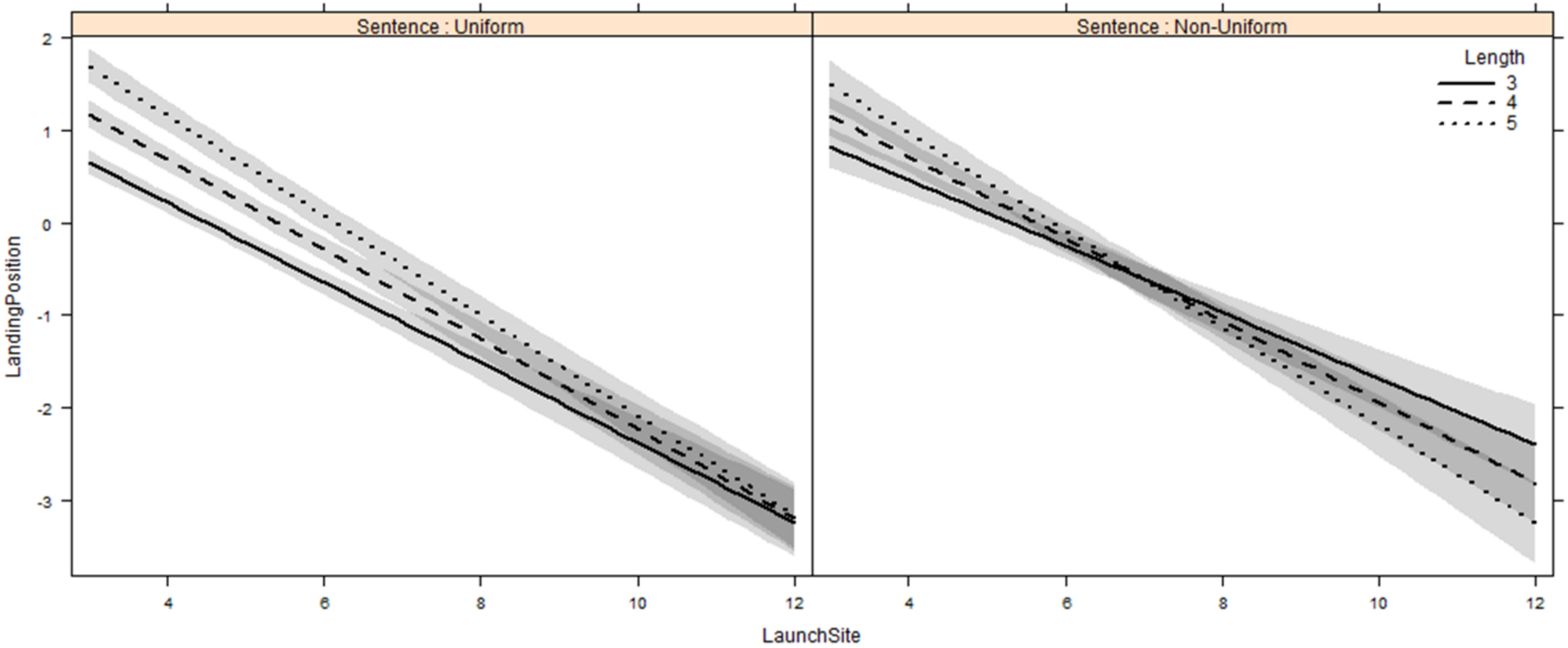
Note. Significant terms are presented in bold. LP = landing position; SU= effect of moving from uniform to non-uniform sentences; WL= word length; LS= launch site.

There were significant main effects of word length and launch site (see Table 4.3 for model output, and Figure 4.1 for estimated effects). As the launch site distance decreased, fixations landed further into a word, thus replicating McConkie et al.'s original observation of the effect of saccade launch site on fixation landing positions. Word length also had a significant main effect, such that the eyes would land further into a long word.

In addition to these main effects, the model also revealed several significant interactions. Sentence uniformity interacted with launch site in a two-way interaction, and with both word length and launch site in a significant three-way interaction. Clearly the word length uniformity within our sentences had an effect on the systematic range error, with it modulating the effect of launch site on landing positions. There was also a significant two-way interaction between word length and launch site. The nature of all these interactions can most clearly be seen by examining Figure 4.1. In the case of saccades made from a near launch site (i.e. 3 or 4 characters) in non-uniform sentences, participants overshot the centre of five letter words to the greatest extent and the centre of three letter words to the smallest extent. For saccades made from far launch sites (i.e. 7 or more characters) in the non-uniform sentences participants undershot the centre of five letter words to the greatest extent, and three letter words to the smallest extent. For intermediate launch sites within non-uniform sentences participants neither undershot nor overshot the word centre. Crucially, the model predicted that, in the case of all three word lengths in the non-uniform sentences, participants would accurately fixate the word centre when making saccades of between 5 and 6 characters. To obtain estimates of the point at which participants would accurately fixate the word centre we used the Effects library (Fox, Weisberg, Friendly, & Hong, 2015) in R, and examined the output for the launch site that predicted a landing position of 0. The model predicted that participants would land directly in the word centre (i.e. a landing position of 0) given launch sites of 5.29, 5.63, and 5.85 characters for three, four, and five letter words, respectively. This suggests that within our non-uniform sentences the preferred saccade length was between five and six characters, with any differences on the basis of word length being reasonably small. The mean of these three preferred saccade lengths was 5.59 characters. While this preferred saccade length differs quantitatively from that obtained by McConkie et al. (1988), most likely due to the lower mean word length in our non-uniform sentences, our results do qualitatively replicate the original finding of a similar preferred saccade length for words of different lengths in normal English sentences.

A different pattern of results emerged for saccades made within our uniform sentences. Similarly to the non-uniform sentences, participants overshot the centre of a five letter word from near launch sites to a greater extent than the centre of three letter words. However, while in the non-uniform sentences the effect of word length quickly reversed with an increase in launch site, such that people undershot the centre of a five letter word to a greater extent than the centre of a three letter word from a far launch site, this was not the case across the uniform sentences. What is more interesting to our investigation of the adaptability of the preferred saccade length is the launch site from which participants would on average land perfectly in the centre of words of each length, and the way in which this varied across our three different types of uniform sentences. To estimate this we again used the Effects package in R. Participants would land directly in the centre of three, four, and five letter words from launch sites of 4.52, 5.41, and 6.14 characters, respectively. Thus, while the preferred saccade length was relatively similar for words of different lengths in the non-uniform sentences, it systematically varied across our three uniform sentence types. This is highly suggestive of adaptation in the preferred saccade length dependent on sentential word length context.

Figure 4.1. *Fixed effects estimates from the linear mixed effect model for initial fixation landing position as a function of sentence uniformity, word length, and saccade launch site. Both landing position and launch site were calculated relative to the word centre. 95% confidence bands are presented around the lines of estimated effects.*



In addition to the main analysis presented above, we briefly considered whether our adaptation effect increased throughout an experimental block, or whether it was rapidly established, remaining constant across the majority of trials. To investigate this we constructed linear mixed effects models with trial index included as an additional predictor. We added this first as a main effect, and then as part of interactions with other factors. Model comparisons were conducted with each increase in model complexity. While a model including only a main effect of trial index significantly improved the fit of the model to our data [$\chi 2(1) = 1995$, $p < .001$], allowing trial index to interact with other factors in two-, three-, or four-way interactions did not further improve our model. Furthermore, while the inclusion of trial improved the fit of our model, this effect was not significant within the model. The lack of interaction between trial index and other factors suggests that our adaptation effects were very rapidly, if not immediately, established.

As mentioned above there were small differences in the frequencies of the words in our different sentence types. To assess whether this had any influence on our results we included frequency in our model as a main effect, and in several interactions. The most complex model which led to an improvement in the fit to the data included an interaction between frequency and word length. All of the effects from our original model remained significant in this model, alongside a main effect of frequency and the interaction between frequency and word length. We obtained estimates of the preferred saccade length from this new model, and obtained very similar results as from our original model; three, four, and five letter words in uniform sentences had preferred saccade lengths of 4.46, 5.45, and 6.25, whereas in non-uniform sentences the preferred saccade lengths were 5.14, 5.6, and 5.9. Thus, our adaptation effects were still present when accounting for word frequency.

4.3.2.3. Fixation Landing Position Distributions. In addition to our linear mixed effects analyses, we conducted a further analysis which was more in line with that presented by McConkie et al. (1988). To conduct this analysis, we transformed our continuous landing position data into discrete values, such that a fixation was categorised as landing on a certain letter within a word (or the space preceding it). Figure 4.2 shows the distribution of initial fixations on each letter of a word, as a function of word length, sentence uniformity, and saccade launch site. Saccade launch sites of 1 to 7 characters from the left of the space preceding the fixated word were used, as in McConkie et al. A landing position of 0 represents the space preceding the word, a landing position of 1 the first letter of the word, and so on.

The standard launch site effect can be seen in Figure 4.2, with a greater proportion of fixations being made on the earlier characters of a word as the distance of the launch site increased relative to the start of the targeted word. This is the case for all three word lengths, in both uniform and non-uniform sentences. Thus, we replicated McConkie et al.'s original findings.

It could be argued that differences should be present for words appearing in uniform relative to non-uniform sentences in these landing position distributions. The logic behind this argument is that the landing position of a fixation within a word should be influenced by the distance from which the saccade was launched, and the extent to which this launch site distance deviates from the preferred saccade length. Furthermore, our linear mixed effects model demonstrated that the preferred saccade length systematically varied between our different sentence types, such that a word of a specific length is subject to a different preferred saccade length when it appears in a uniform relative to non-uniform sentence. Due to this, the distribution of landing positions from each launch site should differ depending upon sentence uniformity, since the deviation of the launch site from the preferred saccade length will differ under these two different conditions. Despite this, there is very little evidence in our landing site distributions for these differences.

While this seems to run contrary to the idea that we observed adaptation in the preferred saccade length, when taking into account both the typical magnitude of the systematic range error and the size of the adaptation in the preferred saccade length, this outcome is not particularly surprising. As discussed above, the range error leads to 0.40 characters of systematic error for each character that the planned saccade length deviates from the preferred saccade length. We can gain a rough estimate of how much the preferred saccade length changed for a word of a certain length by taking the estimate of the preferred saccade length for that word in a uniform sentence from the preferred saccade length for a word of the same length in the non-uniform sentence. In the case of four letter words this equates to a change of 0.22 characters, and for five letter words a change of 0.29 characters. From any given launch site these changes in the preferred saccade length will lead to differences of 0.09 and 0.12 characters of systematic error, respectively. It is not surprising that a shift of such a small magnitude does not lead to a large shift in landing site distributions, especially considering that such subtle effects could have been lost when transforming our continuous landing position data into discrete, whole character values.

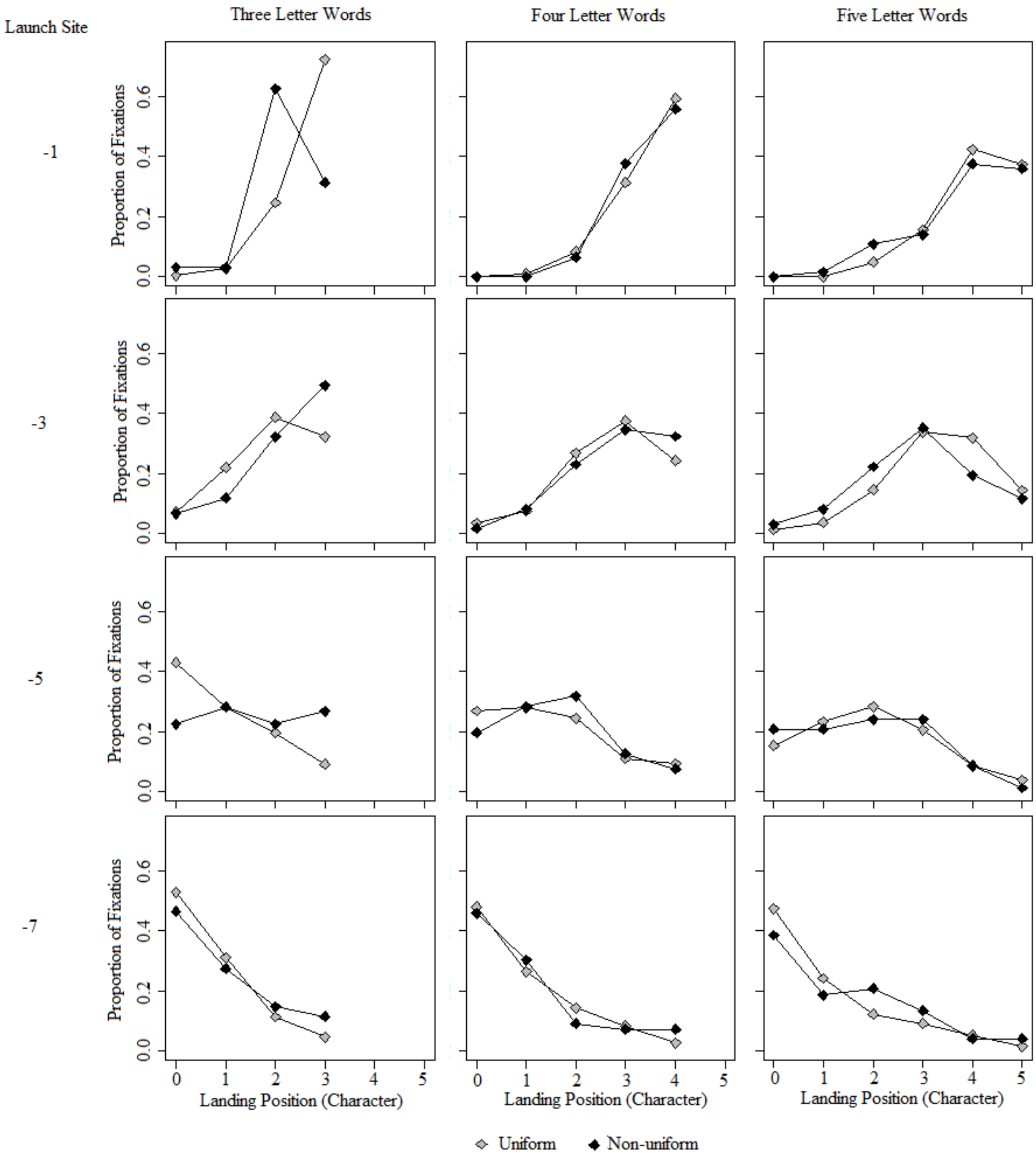


Figure 4.2. Fixation landing position distributions as a function of word length, sentence uniformity, and launch site. Both launch site and fixation landing position were calculated relative to the space on the left of the word. Launch sites between -1 to -7 were chosen in order for equivalence with McConkie et al. (1988).

4.3.2.4. Relative Proportions of Saccades Under and Overshooting. As a final step in our analysis, we also considered the relative frequency with which participants under and overshot the central character of a word from any given launch site in each condition. To examine this issue we categorized our landing position data in terms of whether a fixation landed within a character sized region in the centre of a word, in the area to the left of this central region (i.e. undershoot), or in the area to the right of this central region (i.e. overshoot), and plotted this as a function of word length, sentence uniformity, and launch site (see Figure 4.3). While this categorisation is a coarser measure of fixation landing positions compared to the more typical distributions presented in Figure 4.2, it allows us to observe any systematic variation in the saccadic range error and preferred saccade length more clearly. By considering the launch sites between which participants make a transition from predominantly undershooting the centre of a word to predominantly overshooting that point, it is possible to obtain further confirmation of the variations in preferred saccade length across different sentence types obtained from our linear mixed effects models.

Predictions of the preferred saccade length derived from our linear mixed effects models are as follows. Recall that the linear mixed effects models predicted preferred saccade lengths of 5.29, 5.63, and 5.85 characters for words in these sentences. As such, it seems reasonable to predict that participants should predominantly overshoot for launch sites nearer than these values (i.e. five characters or less) and undershoot for launch sites further away (i.e. six characters or more). As such, readers should transition from over- to undershooting between launch sites of 5 and 6 characters. In contrast, the transition point will vary across the three different uniform sentence types, with the predictions of the preferred saccade length from our linear mixed effects models being 4.52, 5.41, and 6.14 characters for three, four, and five letter words. For each uniform sentence type we should expect the transition point to vary around these values, and as such it should occur between launch sites of 4 and 5, 5 and 6, and 6 and 7 characters for uniform sentences of 3, 4, and 5 letters, respectively.

For the most part, we observed the exact pattern of effects we predicted, with the transition from predominantly overshooting to undershooting occurring at the point we predicted for all three word lengths in uniform sentences, and four and five letter words in non-uniform sentences. In the case of three letter words in the non-uniform sentences the transition actually occurred slightly later, with readers still overshooting very slightly more often than undershooting from a launch site of six. The fact that the transition point was very similar for different word lengths in non-uniform sentences, but increased systematically for different word lengths in the uniform sentences is entirely consistent with the findings from our linear mixed effects model.

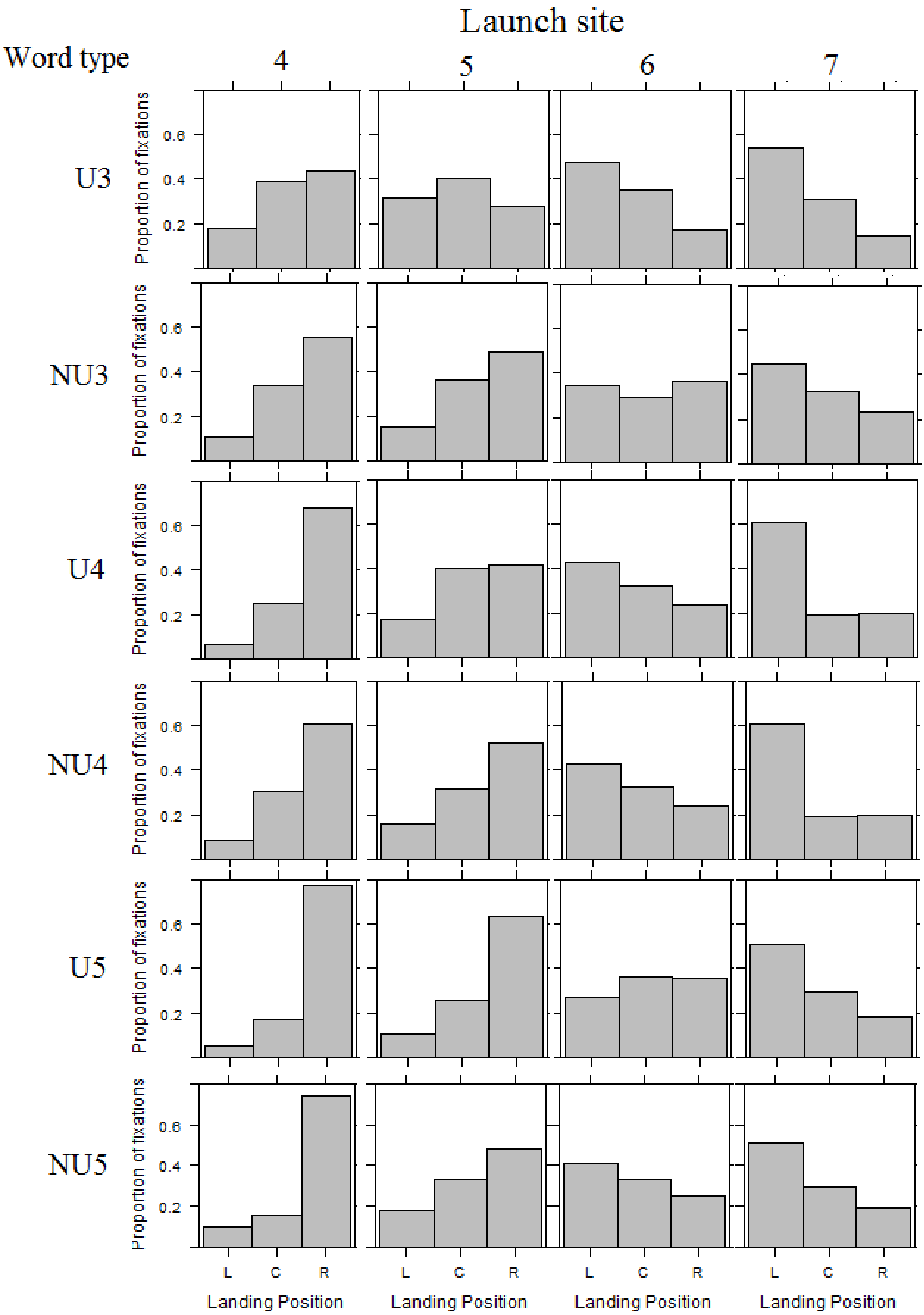


Figure 4.3. Fixation landing positions as a function of word length, sentence uniformity, and saccade launch site. Launch sites are relative to the central point of the fixated word. Launch sites of four to seven characters are presented due to our predictions of when readers would transition from over to undershooting the centre of a word.

4.3.2.5. Overall Saccadic Accuracy. In addition to considering whether our manipulation led to an adaptation of the preferred saccade length, we examined whether this resulted in more accurate overall saccadic targeting. We did this by considering the mean landing position and proportion of initial fixations made on the central character of words of different length in uniform relative to non-uniform sentences. The mean landing position tended to be slightly to the left of the word centre across all conditions, with very little difference between words appearing in a uniform versus non-uniform sentence (see Table 4.4). Furthermore, the proportion of fixations landing on a word's central character also varied very little between uniform and non-uniform sentences (see Table 4.4). Thus, it seems that an adaptation in the preferred saccade length does not translate into more accurate saccadic targeting overall. This may well have occurred as a result of random motor error remaining relatively constant across conditions, and thus leading to a similar distribution of landing positions around the preferred viewing location. Furthermore, these landing positions would then function as the next launch site, which would presumably be non-optimal for the preferred saccade length within the sentence being read.

Table 4.4

Fixation Landing Position and Skipping Probabilities as a Function of Word Length, Sentence Uniformity, and the Difference (D) Between each Word Length in Uniform and Non-uniform Sentences.

	Three letter words			Four letter words			Five letter words		
	U	UN	D	U	UN	D	U	NU	D
LP	-0.45	-0.40	-0.05	-0.41	-0.33	-0.08	-0.34	-0.54	0.20
CLP	0.27	0.25	0.02	0.23	0.23	0.00	0.23	0.22	0.01
SP	0.26(0.43)	0.40(0.49)	-0.14	0.20(0.40)	0.22(0.42)	-0.02	0.13(0.33)	0.11(0.32)	0.02

Note. U= uniform; NU= non-uniform; D= difference; LP= mean landing position relative to the word centre; CLP= proportion of fixations landing in the word centre; SP= skipping probability.

In summary our manipulation of sentence uniformity and the length of words within these sentences had significant effects upon saccadic targeting. First, saccade amplitudes were affected by the length of words within a sentence, such that longer saccades were made for sentences comprised of longer words. Furthermore, variation in saccadic targeting led to an adaptation of the systematic range error, as demonstrated across three different approaches to analyzing the landing position data. Participants were more attuned towards making short saccades in the uniform sentences of three letter words, and more attuned towards making long saccades in the uniform sentences of five letter words, with the preferred saccade length increasing across the three different uniform sentence types. This pattern of effects was very consistent across three different approaches we adopted in our analysis.

4.3.3. Fixation Probabilities and Word Skipping

In order to examine the effect of our manipulation on word skipping we calculated the proportion of words fixated in each sentence (see Table 4.1). Clear word length effects were observed, such that a lower proportion of words were directly fixated in uniform sentences consisting of short words as opposed to long words. However, skipping was no greater in the uniform sentences comprised of three letter words than non-uniform sentences, despite the non-uniform sentences on average containing longer words (*m*=3.94). Thus, while word skipping increased when there were more short words in a sentence, the effect is limited.

Table 4.5

Linear Mixed Model Analyses for Skipping Probability.

	SP		
Model	*b*	*SE*	*z*
Intercept	-1.5	0.25	-5.86
SU	**1.5**	**0.29**	**5.14**
WL	**-0.49**	**0.04**	**-12.68**
Frequency	**0.17**	**0.01**	**15.45**
SU*WL	**-0.33**	**0.07**	**-4.48**

Note. Significant terms are presented in bold. SP= skipping probability; SU= effect of moving from uniform to non-uniform sentences; WL= word length; LP = landing position.

We also examined skipping of individual words as a function of their length, whether they appeared in uniform or non-uniform sentences, and their frequency. As a reminder, we expected a large decrease in the skipping of three letter words in uniform relative to non-uniform sentences, and a slight increase in the skipping of five letter words in uniform relative to non-uniform sentences. We also predicted main effects of length and frequency, two variables that are established as important predictors of word skipping (see Brysbaert, Drieghe, & Vitu, 2005). Highly frequent words and short words were skipped significantly more often than less frequent words and longer words (see Table 4.4 for means and Table 4.5 for LME output). There was also a significant effect of sentence uniformity, which interacted with word length, such that three and four letter words were skipped more often while five letter words were skipped less often in non-uniform sentences than in uniform sentences.

4.4.1. Discussion

In the current study we examined the effect of within sentence word length uniformity on eye movement control during reading, with a particular focus on how this manipulation affects saccadic targeting. Several novel effects were observed. First, there was clear evidence of an adaptation of the systematic range error first observed by McConkie et al. (1988), with the preferred saccade length within a sentence being systematically modulated by the length of words in each sentence type. Furthermore, our manipulation affected word skipping, such that the probability of skipping short words dramatically decreased and the probability of skipping long words slightly increased in uniform relative to non-uniform sentences. Before discussing potential implications, it is worth addressing whether our findings may have been driven by factors other than word length and uniformity.

As we reported in our materials section, two confounds emerged in our stimuli. Potentially, these confounds could have produced our patterns of effects, rather than our word length and uniformity manipulations. The first confound was that the three letter words in the uniform sentences were less frequent than the three letter words in the non-uniform sentences. The second confound was that our uniform sentences were rated as being less natural than our non-uniform sentences. To deal with the first of these issues, we included word frequency as a variable in our linear mixed models, and still found effects of our uniformity manipulation. We, therefore, have evidence that our effects were not caused by the reduced frequency of the words comprising some of our uniform sentences. Also, we note that as frequency decreased, then we would have expected skipping rates to also reduce, and as a consequence of this, landing positions to shift to a slightly earlier position in a word. Given this, the frequency differences between our three letter words in uniform sentences and non-uniform sentences could have contributed to the pattern of results we observed. However, the more pertinent point to make is that frequency decreased as word length increased across our uniform sentences, such that the words in uniform sentences formed from three letter words were more frequent than the words in uniform sentences of four letter words, and these were more frequent than the words in uniform sentence of five letter words. On this basis, we would have predicted that saccade lengths would have been shorter for the longer uniform words. This was clearly not the case, and our results are therefore inconsistent with a frequency based explanation.

In terms of the differences in naturalness, it is worth considering what pattern of effects would have been observed had our results been driven by this factor. As sentence naturalness decreases, reading should become more difficult, resulting in increased reading times (though as stipulated above, we will not focus on reading times), decreased word skipping, and shorter saccades. Thus, such effects should occur for all three word lengths in the uniform relative to non-uniform sentences. Furthermore, any decrease should have been smallest for three letter words due to these uniform sentences being closer in naturalness (m= 3.69) to the non-uniform sentences (m= 3.87), than the uniform sentences of either four (m=3.45) or five letter words (m=3.47). However, in terms of both mean and preferred saccade lengths this was clearly not what we obtained. There was a large decrease in the saccade lengths for three letter words in uniform sentences relative to non-uniform sentences (0.77 characters and 0.93 characters in preferred and mean saccade length, respectively), a smaller decrease for four letter words (0.22 characters and 0.35 characters), and an increase for five letters words (0.29 characters 0.38 characters). Thus, differences in naturalness across our uniform relative to non-uniform conditions could not account for our effects. On the basis of these arguments, we believe it appropriate to discount frequency and naturalness as major contributing factors to our findings.

4.4.2. Saccadic Targeting and the Systematic Range Error

The observation of the systematic saccadic range error by McConkie et al. (1988) has had long lasting implications for the understanding of eye movement control during reading. It is an important parameter in explaining the landing positions of fixations, and is thus implemented in both the E-Z Reader and SWIFT models of eye movement control. Currently both models incorporate the preferred saccade length as a fixed parameter, optimized for saccades traversing the mean distance between the centres of two words in English (E-Z Reader) and German (SWIFT).

Our results show that there was clear adaptation in saccadic targeting across our uniform sentences. This effect can be observed most clearly by considering the preferred saccade length. The estimates from our linear mixed models demonstrated systematic variation on the basis of the mean word length within the sentence being read. The preferred saccade length increased across the three different uniform sentence types, with preferred saccade lengths of 4.52, 5.41, and 6.14 for uniform sentences of three, four, and five letter words. In contrast, any variations between these three word lengths in the non-uniform sentences were more modest, with preferred saccade lengths of 5.29, 5.63, and

5.85. This represents a clear and systematic modulation of the saccadic range error by the current word length context. Our results replicated McConkie et al.'s original finding that landing positions within words are driven by the saccade launch site, in addition to extending those findings by demonstrating an adaptation of the preferred saccade length under different word length contexts. This adaptive modulation of the preferred saccade length may be something that could be considered in future implementations of models of eye movement control.

It is worth considering the launch site from which on average the preferred saccade length was launched from for our three uniform sentence types. Figure 4.4 shows an example of part of each uniform sentence type, with the preferred saccade length to the word centre plotted above. The quantitative adaptation in the preferred saccade length essentially allowed participants to perform the same qualitative behavior in all three uniform sentence types, with participants adapting to move from the preferred viewing location of one word to the centre of the next. This aspect of our results supports our hypothesis that participants would adapt their preferred saccade length alongside the distance between the centres of two words in a sentence.

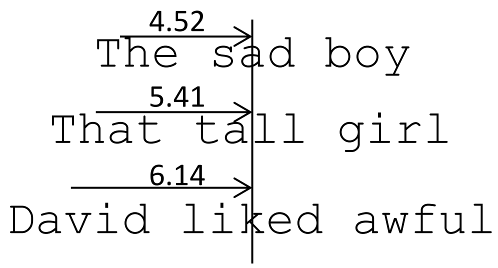


Figure 4.4. The preferred saccade lengths obtained from our linear mixed effects models mapped onto each uniform sentence type. The vertical black line bisects the centre of a word in each sentence type, with the arrow above showing the origin of the preferred saccade length towards this location.

Our results contribute to a literature suggesting a high degree of adaptability during reading. Kaakinen and Hyönä (2014) demonstrated that the extent of the effective visual span during reading is modulated by task demands, and the relevance of a particular piece of text to these task demands. Schotter, Bicknell, Howard, Levy, and Rayner (2014) have shown that the extent to which word frequency and predictability influence fixation durations can be independently modulated by different proofreading tasks, in order to perform these tasks as efficiently as possible. The results from these two studies suggest that the deployment of attention and nature of linguistic processing can adapt during reading based on explicit task demands. Our findings demonstrate that this adaptation extends all the way down to the parameters influencing the execution of saccades. Furthermore, the adaptation in our study occurred despite participants not being conscious of the manipulation, suggesting that these effects are highly automatic.

One surprising aspect of our findings is the time course across with which adaptation occurred. An additional analysis of our data showed that allowing trial number to interact with the other variables in our model for landing position did not improve the fit of the model. This could suggest that our effects were instantaneous, though this would be surprising, since presumably at least some practice is required for adaptation to occur. Due to this, we believe that adaptation very likely occurred across a small number of trials, and quickly plateaued (i.e. adaptation was very rapid, rather than instantaneous). It is clear that future work is needed to examine the time course of our effects.

Our adaptation effects may be difficult to account for within current theories of the systematic range error. One such theory describes the systematic range error in terms of Bayesian decision theory (Engbert & Krügel, 2010; Krügel & Engbert, 2014). According to this approach the oculomotor system makes use of prior knowledge about the distribution of the distance of previous saccade targets from the saccade launch site when estimating the location of the next saccade target. The prior in this theory is essentially equivalent to the preferred saccade length. This prior is combined with a sensory estimate of where the centre of the next word is to yield an optimal estimate of the distance of the saccade target. This combination results in the systematic range error. When a target (and thus its sensory estimate) is far away relative to the prior distribution then the Bayesian combination results in the optimal estimate being nearer than it actually is, and thus saccadic undershoot. When a target is near relative to the prior then the optimal estimate will lead to a saccadic overshoot.

While our findings do not directly contradict the assumptions of this theory, they do suggest that further specification may be required regarding the establishment of the prior. Currently, this theory does not specify the time course across which prior information is taken into account. The simplest assumption would be that the prior is determined over an extended period of reading experience, rather than constantly being reset due to slight changes in mean word length. Under this assumption it is unlikely that this theory can explain our findings. In our study each condition consisted of only thirty trials, and analyses taking trial index into account suggested that the adaptation occurred across the early trials, rather than gradually across all thirty trials. Even if the prior was instantaneously updated with the new information, it seems unlikely that enough information would have accumulated in the early trials of a block to alter a prior established over years of reading experience. Rather, we would expect to see a gradual adaptation across the entire block, as more and more information is integrated into the existing prior. Under such circumstances it is unclear how a substantial adaptation of the preferred saccade length would occur, since the optimal estimate of a saccade target would be biased to a considerable extent by the same prior. In order for this approach to explain our findings it may be necessary for the prior to be based on only a recent number of trials, or to allow multiple priors, which readers switch between dependent upon the current distribution of target distances. Berniker, Voss, and Kording (2010) recently demonstrated that it is possible for people to rapidly estimate multiple Bayesian priors for the same perceptual task. In this study participants had to determine the location of a target on the basis of a noisy perceptual cue and their estimate of the prior distribution of the target's location. The mean of the target's position (and thus the prior) was varied between one of two values, with the current distribution switching throughout the experiment. Participants performed in a manner congruent with Bayesian decision theory, and their performance suggested that they estimated the mean of each prior within ten trials. Furthermore, once both priors had been learnt, participants were able to detect a switch in target distributions within two trials, and begin using the appropriate prior. Thus, it is certainly possible, within the context of Bayesian decision theory, for participants to learn and make use of multiple priors. Given our findings, this is something that should be considered within Engbert and Krügel's (2010) model.

The fact that the preferred saccade length adapts should perhaps be unsurprising, considering that readers will regularly encounter texts differing in terms of the mean of and variability in word length. Within a language, mean word length will vary across texts produced with different purposes in mind. For example, a scientific paper about particle physics should contain longer words than an article in a tabloid newspaper, and the words in a tabloid newspaper may be longer than in something that is read to a child as a bedtime story. It is plausible that all three types of writing will be read by a single person in a day; as such a flexible and rapidly adaptable preferred saccade length is necessary in order for readers to efficiently process a range of texts written for different purposes.

Adaptation may also be necessary for multilingual readers. Different languages vary in the length of the words used. For example, dramatic differences exist between Chinese, English, and Finnish, both in terms of the variations in word length and the mean word length. Liversedge et al. (2016) translated passages of text between these three languages, finding dramatic differences in the word lengths in each language. In Chinese, English, and Finnish the mean word length was 1.55 Chinese characters, 5.63, and 8.32 alphabetic characters with standard deviations of 0.20, 0.80, and 1.44. Furthermore, the saccade length of participants reading in each language varied, with mean rightward saccades of 3.19, 8.53, and 9.35 characters. Clearly, different preferred saccade lengths would be appropriate for these different situations, and flexibility would be vital for bi- or trilingual readers of these languages. It might even be the case that the adoption of a preferred saccade length is more appropriate in a language such as Chinese with relatively little variation in word length, relative to a language like Finnish with highly variable word length.

Further research is required to fully understand the nature of the adaptation observed in the current study. Some issues that need to be addressed include the time course of the adaptive process, the extent to which adaptation occurs when mean word length varies between sentences but is not uniform within sentences, and the extent to which our effects were driven by word length in letters as opposed to spatial extent (see Yao-N'Dré, Castet, & Vitu, 2014). Investigation of these issues will allow the development of a more comprehensive theory of the systematic range error, as well as a greater understanding of the manner in which the saccadic targeting system adapts to the demands of a specific task.

4.4.3. Skipping Probabilities

We observed clear word length effects on skipping. The proportion of words fixated was lower in the uniform sentences of short words than of long words. Furthermore, within the non-uniform sentences, short words were skipped more than long words. There was also an effect of uniformity. The proportion of words fixated was as low in the non-uniform sentences as in the uniform sentences of three letter words, despite the latter condition consisting entirely of words that should be highly likely to be skipped. Furthermore, skipping of three and four letter words was lower in the uniform than non-uniform sentences, while skipping of five letter words was actually slightly higher.

It should perhaps not be surprising that the skipping of three letter words decreased so dramatically in uniform relative to non-uniform sentences. While three letter words were skipped 40% of the time in non-uniform sentences the skipping rate in these sentences across all word lengths was 20%. A skipping rate of 40% across an entire sentence may leave too few fixations to visually encode information. Thus, it is likely that some of the decrease in skipping of three letter words was due to a need to encode visual information from multiple points in the sentence. While this certainly explains a large amount of our observed effect, it is worth considering whether other factors proposed earlier may also have contributed.

One factor was that readers might use the relative length of a word in a sentence to make inferences about its informational value, and whether it requires a direct fixation. In non-uniform sentences this may contribute to the greater skipping of short words (i.e. three and four letters) relative to long words, since short words are likely to be less informative. When this cue is removed in the uniform sentences, the proportion of skipping driven by this source of information would no longer occur, explaining some of the change in skipping rates. Furthermore, a five letter word may be more likely to receive a direct fixation in non-uniform sentences since they are assumed to be of higher informational value than the shorter words, whereas in a sentence of only five letter words there is no reason to assume that any of these words would be of higher informational value than any other. As such, there would not be a tendency to fixate more of these five letter words under the assumption that they may be of more informational value than other words in the sentence. It should be noted that if this argument is correct then this would constitute a demonstration of the adaptability of the cognitive system during reading. Further research is required to evaluate whether this account is correct, or if the effect of relative word length can be explained solely by the need to fixate a minimum proportion of words in a sentence.

A decrease in both accidental and failed skipping due to the saccadic range error may also have contributed to the alteration in skipping behavior. Often a three letter word will be skipped after an intended short saccade as a result of saccadic overshoot, while a five letter word may sometimes be accidentally fixated after a long saccade targeted towards the following word, as a result of saccadic undershoot (see Nuthmann, Engbert, & Kliegl, 2005). In the uniform sentences of three letter words the preferred saccade length was shorter than in the non-uniform sentences, while in the uniform sentences of five letter words it was greater. As such, a saccade targeted towards a three letter word from a near launch site was less likely to overshoot in these uniform sentences than the non-uniform sentences, while saccades programmed to skip a five letter word from any given launch site would be less prone to undershooting in the uniform sentences. This may have contributed to the decreased skipping of three letter words and increased skipping of five letter words in uniform sentences. It is difficult to be certain of the exact extent to which the adaptation in the preferred saccade length may have affected skipping rate due to mislocated fixations, but it seems reasonable to assume that it had some influence.

4.4.4. Conclusion

In closing, the current paper examined the effect of within sentence word length uniformity on eye movement control during reading. The most striking effect of this manipulation was on the saccadic targeting system, with the preferred saccade length which determines the systematic range error adapting contingent upon the length of the words within a sentence. This effect has implications for models of eye movement control, and theoretical accounts of the systematic range error. We also observed effects of our manipulation on the skipping probabilities of words within the sentences. We argued that the effect on skipping was due to a combination of needing to fixate a minimum number of words in a sentence, as well as relative word length information and the adaptation of the preferred saccade length. Our findings suggest a high degree of adaptability in the saccadic targeting system during reading.

Chapter 5

General Discussion

5.1. Basic Overview of Research and Findings

The findings of the three empirical papers presented in this thesis have implications for our understanding of important theoretical issues in the area of oculomotor control during reading. While all three of these empirical papers dealt with relatively diverse issues, the findings of all of them share several implications for the theoretical understanding of eye movement control during reading, and the computational models stemming from this understanding.

It is briefly worth recapitulating what exactly was investigated, what was found, and what was concluded in each paper presented in this thesis. In Chapter 2 we investigated the processing of a multi-word unit, looking at the specific example of a spaced compound such as *teddy bear*. We were interested in seeing if the constituent words of this multi-word unit were lexically processed as two separate words, or in a manner similar to a single long word. This was investigated by using the boundary paradigm to present participants with various different previews of these constituent words. The spaced compound was placed immediately after an invisible boundary, such that *teddy* was word *n*+1 and *bear* was word *n*+2. The preview of each of these words was manipulated to either be correct or an orthographically illegal non-word (e.g. *fohbg* for *teddy* and *hocu* for *bear*). Participants were shown either a correct preview of both words, an illegal preview of both words, a correct preview of word *n*+1 with an illegal preview of word *n*+2, or an illegal preview of word *n*+1 with a correct preview of word *n*+2. The findings of prior research suggested that we should not observe an effect of the *n*+2 preview type given the processing difficulty of our word *n*+1 (which was comparatively high). However, we did observe an effect of the *n*+2 preview type, but only when there was a correct preview of word *n*+1, which was the first constituent of the spaced compound. We took this as evidence that spaced compounds have a single lexical representation, and that upon lexical activation of the first constituent, processing expands to both constituents in parallel. It should also be noted that in this study our large parafoveal preview manipulation had no effect on fixation times on word *n*, an issue which I will return to later.

In Chapter 3 we investigated whether orthographic information from word *n*+1 in the parafovea is able to influence the processing of word *n*+2, and vice versa. This question was addressed through the use of the boundary paradigm. We placed the boundary before a three letter word (word *n*+1) followed by a three or four letter word (word *n*+2). Participants were presented with either a correct preview of these two words (e.g. *hot pan*), a preview in which two letters had been transposed between the words (e.g. *hop tan*), or a preview in which the same two letters had been replaced with alternative letters (e.g. *hob fan*). We hypothesised that if orthographic information from both of these words was fed into the lexicon in parallel then we would observe faster reading times on our target words when participants received the correct orthographic information as part of the wrong word (i.e. in the transposition condition) than when this information was entirely absent (i.e. in the substitution condition). However, while we observed significant identity preview effects in this study, minimal differences were present between the transposition and substitution condition. This was the case in both Experiment 1 when the transposition was made between the final and first character of word *n*+1 and word *n*+2 (e.g. *hop tan* as a preview of *hot pan*) and in Experiment 2 when the transposition maintained a letter's within word position (e.g. *pit hop* as a preview of *hit pop*). From this data, we concluded that participants do not process orthographic information from word *n*+1 and word *n*+2 in parallel.

Finally, in Chapter 4 we investigated the extent to which the saccadic targeting system is able to adapt to changes in both the mean and variability of word length within a set of sentences. We investigated this question by presenting readers with four sets of thirty sentences; one set was made exclusively from three letter words, one set was made exclusively from four letter words, one set was made exclusively from five letter words, and the final set was made of a combination of three, four, and five letter words. We undertook a thorough analysis of this data set to assess whether participants changed their approach to saccadic targeting depending on the word length condition. Specifically, we investigated whether the preferred saccade length at which participants are most likely to accurately fixate a target word's centre varies depending on the current word length context. We found that this was the case, with readers being most accurate at making short saccades in the sentences of three letter words, and most accurate at making longer saccades in the sentences of five letter words. Furthermore, the regularity with which participants would skip a three letter word was reduced in the uniform sentences of three letter words compared to the non-uniform sentences. We took both of these findings to be evidence that the saccadic targeting system is able to adapt to novel word length conditions, both in terms of the preferred saccade length and which words are selected as a saccade target.

5.2. Assessing E-Z Reader and SWIFT

Before discussing the common implications of our findings, I will briefly consider the extent to which each of the computational models outlined in the introductory chapter are able to explain the findings from each study. While E-Z Reader has not yet attempted to simulate our findings, in all likelihood it would be possible to modify its current parameters to do so, without violating its theoretical underpinnings. To begin with, the data we reported in Chapter 2 can be accounted for by E-Z Reader at a theoretical level, although not computationally. According to E-Z Reader, attention would have shifted to word *n*+1 in the parafovea, accounting for our *n*+1 preview benefit; this is something E-Z Reader can simulate easily. What might be less simple for E-Z Reader to simulate is our *n*+2 preview effect. However, our key argument in this paper was that the information from both word *n*+1 and word *n*+2 was processed in parallel, but as part of a larger lexical unit. Since in this case lexical processing was still confined to a single unit, this is not a problem for the theoretical assumptions of E-Z Reader. Rather, it is only the case that our findings cannot be simulated by E-Z Reader simply because it has not been designed to account for lexical processing at such an in-depth level.

The findings of the study discussed in Chapter 3 are entirely in line with E-Z Reader. Again, since attention will typically shift to word *n*+1 prior to a saccade being made towards this word, the identity preview effects we observed are easily explained by this model. The identity preview effects on word *n*+2 can also be accounted for by this model, either as a spillover effect from having processed the incorrect preview of word *n*+1, or due to attention being able to shift to the incorrect preview of word *n*+2 prior to a saccade crossing the boundary. In our study word *n*+1 was always only three letters long, and so the probability of identifying this word prior to crossing the boundary was quite high. Indeed, the skipping probability of this word was around 30%, suggesting that it was regularly identified in the parafovea. Furthermore, since information from word *n*+2 should not be processed at the same time as word *n*+1 it is no surprise, from the perspective of this model, that orthographic information from word *n*+2 did not influence the processing of word *n*+1. Indeed, this finding may even lend support to the idea put forth by Reichle, Pollatsek, and Rayner (2006) that orthographic parafoveal-on-foveal effects are due to unusual visual information, rather than two words being processed in parallel.

Finally, the main finding from Chapter 4 cannot currently be explained by E-Z Reader, although once again a simple parameter change would be sufficient. In this case, it would simply be a matter of allowing the preferred saccade length of 7 characters to become flexible towards the characteristics of the text currently being processed. Once again, this would in no way violate the theoretical assumptions of E-Z Reader.

SWIFT may actually be able to account for more of our data than E-Z Reader at a computational level; however, I consider our findings to be more difficult for the model to explain at a theoretical level. In terms of Chapter 2, it is not problematic for SWIFT that an *n*+2 preview effect was obtained, dependent upon the orthographic legality of word *n*+1. What is problematic is that we obtained this effect by using spaced compounds as our stimuli, in contrast to prior studies that have failed to observe these *n*+2 preview effects. Given that SWIFT proposes that words within the perceptual span should always be processed in parallel, it is hard to see how it would account for the parafoveal words forming a multi-word unit modulating the extent to which they are processed in parallel. Furthermore, as discussed in Chapter 2, one explanation for the lack of *n*+2 preview effects when word *n*+1 is longer than three letters is that word *n*+2 is pushed too far into low acuity vision to be efficiently processed. The fact that we found that information can be extracted from word *n*+2 further into the parafovea under specific circumstances suggests that it is not visual acuity that prevents more prevalent *n*+2 preview effects, so much as the way in which lexical processing typically occurs.

Again, at a computational level, the data presented in Chapter 3 are not problematic for the SWIFT model. As discussed, this model predicts preview benefits of both word *n*+1 and word *n*+2 as a matter of course, and as such observing faster reading times given an identity preview of both words is in line with the predictions of this model. Furthermore, there is no implemented parameter in SWIFT that would predict an effect of transposing two letters between words in the parafovea rather than substituting them. Indeed, considering that SWIFT operates on the basis of parallel independent channels (see Section 1.3.3) it should certainly not be predicted that orthographic information from one word should influence the processing of another. As such, this null effect is not technically problematic for SWIFT. However, as discussed at length in this chapter, it is not clear how multiple words could plausibly be processed in parallel without information from one word influencing the processing of another word. While the use of parallel independent channels may seem to alleviate such issues by not allowing information from separate channels to interact, this may not necessarily be the case. Presumably, information from each channel must still cause activation in the lexicon at the same time, thus potentially leading to the effects we predicted in this study. It could hypothetically be the case that each channel is independent to the extent that each word in the perceptual span is fed into a separate lexical processing system, but this would seem to create more issues for SWIFT than it would solve. For example, would each of these independent lexicons learn separately? In addition, implementing independent lexicons would compromise SWIFT's ability to explain other phenomena dependent upon all words being processed in the same lexicon (e.g. within sentence inhibitory neigbour priming; see Paterson, Liversedge, & Davis, 2009). As such, our finding may well become problematic if a word identification module was implemented in SWIFT.

The finding that the preferred saccade length can adapt is not currently predicted by SWIFT, with this model implementing a fixed preferred saccade length of 5.4 characters. However, as was the case for E-Z Reader, altering this would be fairly simple, and would not violate any of the theoretical underpinnings of the model.

In summary, neither model is able to completely explain all of our findings. Furthermore, it is not necessarily the case that the model that would currently have the most trouble explaining each finding is the one for which it is most problematic. For example, in the case of Chapter 2, implementing a parameter to explain the processing of a multi-word unit as a larger lexical item would most likely be less problematic for E-Z Reader than SWIFT. Similarly, the data presented in Chapter 3 can currently be explained by both models quite well, and it is difficult to see a way in which these data would become problematic for any future iteration of E-Z Reader. In contrast, were SWIFT to be developed to include a word identification module, it may be the case that it would begin to predict an effect of our inter-word transposition manipulation. Finally, neither model currently allows the preferred saccade length to vary, but it would presumably be equally easy to alter this in both E-Z Reader and SWIFT.

5.3.1. Common Implications

While the implications of these findings are clearly discussed in the relevant empirical chapters, the ways in which these findings relate to each other may be less obvious. I will now consider some of the common implications of the work presented throughout this thesis, both in relation to the models of eye movement control, and more general theoretical issues. To be clear, the points I make below are not necessarily meant as criticisms of the models, so much as suggestions for ideas that may be implemented in a future version of these models that are able to account for an even more comprehensive set of phenomena observed in the reading process.

5.3.2. Parafoveal-on-Foveal Effects

One issue that all three of the papers presented in this thesis relate to is the concept of parafoveal-on-foveal effects. As discussed throughout this thesis parafoveal-on-foveal effects refer to the characteristics of a word in the parafovea influencing the processing of a fixated word, and are often viewed as vital in the debate between serial and parallel processing accounts. These effects can either be driven by a fairly low-level orthographic manipulation (e.g. presenting *pvxforming* instead of *performing*) or a higher level lexical manipulation (i.e. varying the frequency) of the parafoveal word. There are several ways in which these effects are controversial; they are not always reliably observed; they may actually be the consequences of other phenomena such as mislocated fixations and eye-tracker error; and there is not even a clear agreement about the theoretical explanation of these effects if they are real.

The data with the most obvious link to this issue were presented in Chapter 3. In this paper we presented two experiments which aimed to assess the extent to which orthographic information from one parafoveal word can influence the processing of another parafoveal word. As discussed in depth in Chapter 3, the fact that the orthographic characteristics of word *n*+1 have sometimes been found to influence fixation durations on word *n* could be taken as evidence for letter information from two words feeding activation into the lexicon at once, which would constitute clear evidence for the parallel processing of two words. However, it could also simply be the case that unusual visual configurations pop out at a pre-attentive level, thus disrupting processing at a non-lexical level. By demonstrating that orthographic information from one parafoveal word does not influence the processing of another parafoveal word we showed that it may well be the case that previously observed orthographic parafoveal-on-foveal effects may simply be due to the visually salient nature of the manipulation used, rather than parallel lexical processing. This interpretation is very much in line with the explanation of orthographic parafoveal-on-foveal effects proposed by Reichle, Pollatsek, and Rayner (2006) in the context of the E-Z

Reader model, whereby unusual visual information pops out. It is less in line with an interpretation of orthographic parafoveal-on-foveal effects whereby they are caused by the processing of orthographic information from one word interfering with the processing of another word.

The study presented in Chapter 2 was not primarily focussed on investigating parafoveal-on-foveal effects. Nonetheless, our manipulation in this study may have been expected to lead to the observation of these effects, with either one or two relatively long words in the parafovea being replaced with orthographically illegal non-words. This is a far more dramatic manipulation than is used in some studies that have previously observed parafoveal-on-foveal effects due to unusual orthographic information (e.g. Drieghe, Rayner & Pollatsek, 2008; Inhoff, Starr, & Shindler, 2000). Despite this, we did not observe a significant effect of our preview manipulation on word *n* in this study. Thus, our data constituted a failure to replicate controversial effects that are often taken as evidence in favour of parallel lexical processing.

The theoretical advances presented in Chapter 4 relate to these parafoveal-on-foveal effects in a slightly less obvious way. One explanation for parafoveal-on-foveal effects which reduces them to an artefact of another phenomenon in oculomotor control is the mislocated fixations account (Drieghe, Rayner, & Pollatsek, 2008). According to this account, parafoveal-on-foveal effects occur as a consequence of saccadic error causing participants to fixate a word other than the one they had intended to fixate and process. As a result, participants process the original saccade target while fixating on an adjacent word. Viewing times on this mistakenly fixated word are then influenced by the characteristics of the original saccade target in the parafovea, leading to the observation of a parafoveal-on-foveal effect. The primary source of the saccadic error leading to a mislocated fixation is the systematic range error, which was extensively discussed in Chapter 4. As a general rule researchers have treated the preferred saccade length around which the systematic range error revolves as being fixed at around seven characters, at least in the case of English. The data presented in Chapter 4 clearly contradict this assumption, with the preferred saccade length adapting on the basis of mean word length information. This should be taken into account in studies that observe controversial parafoveal-on-foveal effects that could in fact be an artefact of mislocated fixations. Arguably, the plausibility of mislocated fixations explaining parafoveal-on-foveal effects for any given set of data depends upon the preferred saccade length for the text being read, and the extent to which the intended saccade length which seems to be leading to parafoveal-on-foveal effects deviates from this.

5.3.3. The Need for Flexibility

Another theme that emerges from several of our papers is the need for approaching certain parameters relating to oculomotor control in a flexible manner. As discussed throughout, current models of eye movement control disagree about whether words are lexically processed in serial or parallel. This disagreement is treated in a fairly dichotomous manner, in that either all words within the perceptual span should be processed in serial order, or in parallel. However, the data presented in Chapter 2 and Chapter 3 indicates that this is not the case. Rather, these data suggest that the human cognitive system is flexible in whether or not it processes words in serial or parallel. Specifically, the data in Chapter 3 suggests that, as a general rule, readers will process simultaneously available words in serial order as opposed to in parallel. However, the data presented in Chapter 2 demonstrates that there is an exception to this rule; when multiple spatially separated words form a larger lexical unit, the information from both words will be processed in parallel. Future accounts of eye movement control may need to consider that multiple words can either be processed in serial or parallel depending upon whether they form a larger unit. We also made the point in Chapter 3 that in order for a parallel model to reconcile the fact that orthographic information from word *n*+2 does not influence the processing of word *n*+1 (and vice versa) with the fact that orthographic information from word *n*+1 has been found to influence the processing of word *n* it would need a level of flexibility in how words in different parts of the perceptual span are processed; otherwise, proponents of these models would have to concede that either parafoveal-on-foveal effects or word *n*+2 preview effects are not due to the parallel processing of words within the perceptual span.

The data presented in Chapter 4 also illustrates the need for flexibility in models of eye movement control. As discussed in depth earlier, both SWIFT and E-Z Reader implement the systematic range error first investigated by McConkie et al. (1988). This range error depends upon the preferred saccade length, with participants undershooting the centre of a word when they must make a saccade longer than the preferred saccade length, and overshooting it when they must make a saccade longer than the preferred saccade length. Both models assume that the preferred saccade length is fixed at a constant value.

However, our data directly contradicts this assumption, suggesting that both models should allow a greater level of freedom in this parameter than they do.

5.3.4. A Word Processing Module

This final issue only really relates to the studies presented in Chapters 2 and 3. As it stands, models of eye movement control do not include implemented models of word identification. However, in order to explain some of our findings this may be necessary. For example, in Chapter 2 we discussed our findings in terms of lexical processing being extended further into the parafovea upon readers encountering the first constituent of a spaced compound. We proposed that this could be accounted for within an interactive-activation framework, whereby the first constituent of a spaced compound activates the larger lexical representation of the whole unit, and, via feedback activation from this unit, processing is extended further into the parafovea. At a theoretical level a model such as E-Z Reader is perfectly capable of accounting for this sort of phenomena, due to it not violating the key tenet of serial lexical processing. However, how such a process would actually be implemented in the model is less clear, given that it is currently not discussed exactly how words are processed.

An important point of Chapter 3 was that it is difficult to be sure of how exactly word identification would occur in a parallel model such as SWIFT. In this paper we argued that it is difficult to envisage how multiple word forms could be fed into the lexicon in such a way that the processing of one word would not interfere with the processing of another. Having failed to observe evidence of such interference in this study, the question becomes how exactly words are processed in a parallel model, in a way compatible with the findings reported in this thesis.

5.4. Avenues for Future Research

I will now consider several ways in which the ideas and experimental paradigms explored in this thesis can be extended into future research. While many of these ideas will be based exclusively on a single chapter, there are also some interesting avenues for combining the ideas and methods explored across multiple chapters.

In Chapter 2 we made the point that spaced compounds are not the only multi-word unit that may be stored in the lexicon. Rather, a large number of other units exist that may be processed in a similar manner. These units include, but are not limited to, idioms (e.g. *kick the bucket*), binomial word pairs (e.g. *fish and chips*), longer compound terms (e.g. *United States of America*), and names. It is necessary for further research to investigate the processing of these units in order to establish multi-word unit processing as a more general phenomenon. It seems likely that not all units will be processed in the same manner as spaced compounds, and it is important to gain an understanding of exactly which factors are important in causing a multi-word unit to be stored and processed holistically. One factor we are especially interested in investigating further is the influence of the extent to which the first constituent of a multi-word unit is predictive of the rest of the unit. In the mechanism proposed in Chapter 2 this should presumably have a large influence, with the level of constraint strongly influencing the extent to which the early constituents of a multi-word unit will activate the whole unit's form, as opposed to individual words (see also Yu, Cutter, Yan, Bai, Fu, Drieghe, & Liversedge, 2016).

There are also several paradigms we could use to further investigate these issues. For example, it has been shown in prior research that when a single word in the fovea disappears 60ms after direct fixation, reading continues uninhibited (Liversedge, Rayner, White, Vergilino-Perez, Findlay, & Kentridge, 2004). In contrast, when the word to the right of fixation disappears 60ms into a fixation, it is highly disruptive (Rayner, Liversedge, & White, 2006). Presumably, this effect is partly due to a visual representation of the foveal word being encoded within this time frame, whereas the parafoveal word should not be encoded until attention shifts towards it. Thus, the disappearance of the foveal word is less disruptive. An interesting extension of the work presented in Chapter 2 would be to test whether this phenomenon is altered when the foveal and parafoveal word form a spaced compound. If our contention that readers treat these units as single words is correct, then the parafoveal word may be encoded sooner than is typical, and as such its disappearance would be similarly undisruptive to the disappearance of the foveal word.

A further interesting extension of the work in Chapter 2 would be to use the paradigm implemented in Chapter 3. By this I mean that readers could be presented with a correct preview of a spaced compound (e.g. *teddy bear*), a preview in which a letter has been transposed between the two constituent words (e.g. *teddb year*), or a preview in which these two letters have been replaced (e.g. *teddl gear*). Presumably, if these two words are processed at the same time, we would observe a benefit of the transposition condition relative to the substitution condition. There is an interesting parallel here to work conducted in Chinese by Gu and Li (2015). They found that when two characters were parafoveally transposed between words in Chinese there was no benefit relative to a substitution preview, whereas parafoveal transpositions within words did provide a benefit relative to substitution previews. It might be that while we observed no benefit of our between-word transposition, we would if these two words formed a larger unit. Indeed, a large literature exists in alphabetic languages showing that transposition benefits occur within words, suggesting that if spaced compounds really are processed as lexical units, then we may expect similar effects.

Beyond applying the manipulation presented in Chapter 3 to spaced compounds, it would be interesting to see how these effects (or lack thereof) are modulated by the lexical status of spatially unified words. As discussed in Chapter 2, there is a level of debate as to whether even unspaced compounds are processed holistically or compositionally. Furthermore, prior research has shown effects of the lexical status of a word on preview and within-word parafoveal-on-foveal effects (Drieghe, Pollatsek, Juhasz, & Rayner, 2010; Häikiö, Bertram, & Hyönä, 2010). Applying our inter-word transposition manipulation to unspaced compounds which vary on a number of factors (i.e. the frequency of the whole compound and the frequency of the constituent words) could give us further insight as to what characteristics determine whether these units are processed via a direct lookup or compositional route.

Finally, there are a large number of ways in which we can extend the work presented in Chapter 4. Before moving onto work focussing on saccadic targeting I will briefly consider how the stimuli used in this study can be used to explore parafoveal processing. First of all, we can use these stimuli to further investigate the issues discussed in Chapter 3, in terms of how the orthographic information from one word is able to affect the processing of other words within the perceptual span. As discussed in Chapters 1 and 3, Angele, Tran, and Rayner (2013) investigated this issue by presenting a repetition of the fixated word (e.g. *news*) in the parafovea, finding that fixation times on the foveal word were shorter under these conditions than without a repetition. We can extend this manipulation to an extreme point, by replacing every word of the sentence in the parafovea with the currently fixated word, and examining whether this provides any further benefit relative to a single repetition. This is a manipulation which would not be possible in sentences with words of varying length, since it is necessary for the replacement words to subtend the same number of character spaces as the words that will be presented upon direct fixation. We could also make this manipulation less visually salient by splitting the repeated orthographic information across multiple parafoveal words; this could be implemented by replacing the last two letters of word *n*+1 with the last two of word *n*, and the first two letters of word *n*+2 with the first two letters of word *n*. For example, if readers were fixated on the word *girl*, the parafoveal previews of the following words (e.g. *must want*) would be *murl gint*.

Along a similar vein, we could also use our stimuli to examine if readers gain a greater benefit from having previews of an upcoming word at multiple positions in the parafovea. This would be achieved by taking a word towards the end of the sentence and using a gaze-contingent paradigm in order to have this target word presented to the right of the currently fixated word until participants have fixated the actual target word. By examining fixation times on this target word given these multiple previews, relative to a single preview, it would be possible to assess whether preview benefits are entirely positionally specific, or if benefits can build up on the basis of multiple previews in sentence positions other than the target word. If it were the case that benefits did build up over multiple previews, we could then use this paradigm to investigate types of preview effects that are often elusive, such as semantic preview effects from non-synonyms and morphological preview effects; by doing this we can determine if they are elusive due to parafoveal words not typically being processed for this type of information, or if it is simply the case that previews are not typically available for long enough for readers to reach a certain stage of processing.

In addition to using our existing stimuli to investigate issues of parafoveal processing, we can also further investigate the flexibility of saccadic targeting. For example, in Chapter 4 we found that our saccadic adaptation effects did not seem to interact with how far into a block of trials a participant was. This raises the question of whether these effects are in fact instantaneous. This could easily be investigated by randomly presenting our uniform sentences amongst filler items, and examining whether we still observe our original effects. Furthermore, the question arises as to whether our effects would be similar if we were to move beyond the range of word lengths used in our original experiment. All of the word lengths we investigated could typically be identified in a single fixation. It could be that we would observe a different pattern of effects if we investigated longer words that would require multiple fixations. We could even extend this work into other languages, such as those with a larger mean word length (e.g. *Finnish*) or a smaller mean word length (e.g. *Chinese*).

5.5. Final Comments

It is my hope that this final chapter has illustrated the important theoretical contributions of the work presented throughout this thesis. While these papers dealt with relatively diverse topics they all have important implications for our understanding of oculomotor control during reading, especially in terms of parafoveal processing and saccadic targeting. Furthermore, it should be clear from the previous section that these studies can be developed far beyond what has been presented in this thesis, in order to further increase our understanding of oculomotor control during reading.

Appendices

Appendix A: Chapter 2 Stimuli

A list of the sentences used in the study presented in Chapter 2. The constituent words of the spaced compounds are italicized. The non-word preview for each constituent is displayed in brackets.

1. He knew the shaman's fortune telling was just silly *mumbo (umnkc) jumbo (gmvke)* that should be ignored.
2. As the witch threw the rat into the cauldron she yelled *hocus (kaovz) pocus (gaawz)* and began stirring.
3. One gym class involved using the waist and hips to spin a large *hula (bmto) hoop (bcaq)* round the body.
4. The deer triggered a tripwire, and was caught in a nasty *booby (haedp) trap (fvcq)* out in the forest.
5. The lighting manager for the local club set up his flashy *strobe (zlmoko) lights (tfdpsz)* for the rave.
6. She gently stroked from the head to the tail of the white *pussy (qrzzj) cat (ool)* sat on her lap.
7. One of her favourite Christmas puddings is tasty *mince (vfmec) pies (gtoz)* with loads of cream.
8. She saw the man with a cold wipe the snot from his gross *runny (vmvrg) nose (meza)* with a tissue.
9. The Muslim man's meal would be incomplete without some tender *halal (kefcf) meat (wael)* in a sauce.
10. The man living near an airport got annoyed by the noisy *jumbo (ymrdc) jets (yofz)* landing so nearby.
11. The small child gently cuddled his fluffy *teddy (fohbg) bear (hocu)* while trying to get to sleep.
12. The Japanese gardener looked sadly at the dying *bonsai (hcvzcf) tree (lnao)* that she had not watered.

13. He loved to go to the Indian and get a spicy *tikka (ftddo) masala (vozetc)* to eat during the football.
14. As it purred she caringly cuddled the happy *tabby (lchdp) cat (oel)* that was lying in her arms.
15. Since the ground was wet and muddy outside she wore her green *welly (uoftj) boots (declz)* to the park.
16. The old dog lover patted the hairy *cocker (oaabou) spaniel (zqomlot)* that came up to him on the head.
17. She tried to complete a thousand piece *jigsaw (qtpzou) puzzle (qwssta)* during her holiday in Wales.
18. At the zoo the child watched the cute, grey, furry *koala (hcetc) bear (kaov)* slowly climb a tree.
19. He was kept up by his housemate's wet clothes being in the noisy *tumble (lvvdto) dryer (hmjow)* all night.
20. The view of the fairground from the top of the giant, round *ferris (tcwwtz) wheel (ubaot)* was astounding.
21. During hide and seek she hid quietly behind the wooden *Wendy (Nomkj) house (kamzc)* in the play area.
22. At the fair she had fun sliding down and around the bright *helter (kaftan) skelter (zhoffon)* several times.
23. The war was confirmed at a press conference by the Dutch *Prime (Rwtno) Minister (ulmlzfom)* last Tuesday.
24. As he climbed the security fence the thief cut his leg on the sharp *barbed (doudcb) wire (ntmo)* on top.
25. The woman greedily ate the entire pot of honey roast *cashew (aezdcv) nuts (rvlz)* all by herself.
26. Baked beans, sausages, and creamy *mashed (uezkch) potato (gelclc)* made up his daughter's favourite dinner.
27. She cooked her eggs on the electric hob in a metal *frying (lwgtvy) pan (gcv)* full of melted butter.
28. The secretary put the document back into the steel *filing (lfftmq) cabinet (oekfmaf)* in the corner.

29. Since it was cold he put on his hat, scarf, gloves and thick *duffle (bnlltc) coat (aacl)* before leaving.
30. At the Indian restaurant she had a side dish of plain *pilau (yttow) rice (ufec)* with her chicken korma.
31. She spread both the jelly and crunchy, thick *peanut (qaevrl) butter (kmffcv)* onto the toast for her son.
32. On Christmas Eve she loved to heat up and drink spiced *mulled (vwffab) wine (vlro)* before going to bed.
33. Using newspaper and glue in art class she made a giant *papier (gcgfav) mâché (vcako)* model of a hippo.
34. The cocaine smuggler looked nervously at the brown *sniffer (zmfllam) dog (haq)* by the airport gate.
35. I stuck the broken box back together using glue, extra sticky *masking (uczhfvp) tape (leqo)*, and staples.
36. If they were going to swim in it, his toddlers' dirty *paddling (jekktfwj) pool (jaef)* needed a good wash.
37. He breaks the shell, beats the yolk and white, and then makes *scrambled (zoueudtok) eggs (cppz)* on toast.
38. The small coastal town had been flooded and ruined by the giant *tidal (ftbef) wave (nowo)* last August.
39. She had seen many underwater fish over summer while *scuba (zamke) diving (hlmlrq)* in the Maldives.
40. He covered the vegetables he wanted to roast in cheap *olive (attuc) oil (ctf)* and put them in an oven.

Appendix B: Reading Time Measures for Chapter 4.

The following Appendix contains reading time measures for Chapter 4 (which were not the primary focus of our main analyses). In order to examine the effect of our manipulation on reading times at a global level we examined the reading time per word across our different sentence types (see Table B1 for descriptive statistics and Table B2 for inferential statistics). Across the three uniform conditions reading times increased alongside word length, with the words within uniform sentences of three letter words and uniform sentences of five letter words being read significantly faster and slower, respectively, than the words within uniform sentences of four letter words, replicating standard word length effects. Furthermore, the reading time per word in the non-uniform sentences was significantly shorter than in the uniform sentences of four letter words, and numerically smaller than in the uniform sentences of three letter words. Given that we observed word length effects between the uniform conditions, and that the mean length of the words in the non-uniform condition was the same as in the uniform sentences of four letter words, this finding suggests a cost of sentence uniformity on reading times.

Table B1

Mean Reading Time per Word for each Sentence Type. Standard Deviations are Presented in Parentheses.

	U3	U4	U5	NU
TPW	250(73)	269(79)	329(100)	247(66)

Note. TPW= reading time per word. U3, U4, and U5 refer to uniform sentences of three, four, and five letter words respectively, and NU to non-uniform sentences.

Table B2

Linear Mixed Effects Model Output for Reading Time per Word.

	TPW		
Model	*B*	*SE*	*T*
Intercept	5.56	0.05	117.59
U3	**-0.07**	**0.03**	**-2.79**
U5	**0.20**	**0.03**	**6.62**
NU	**-0.08**	**0.03**	**-2.49**

Note. Significant factors are presented in bold. TPW= reading time per word. U3, U4, and U5 refer to uniform sentences of three, four, and five letter words respectively, and NU to non-uniform sentences.

As well as examining global reading times, we calculated several measures of fixation times on individual words within sentences, as a function of word length and sentence uniformity. The fixation time measures were first fixation duration (FFD; the duration of the first fixation on a word), gaze duration (GD; the sum of all fixations on a word from the first fixation until a saccade is made to another word), single fixation duration (SFD; the duration of a fixation when it is the only one made on a word) and total time (TT; the sum of all fixations on a word). These measures were log transformed in order to increase normality and were analysed using linear mixed effects models. Sentence uniformity and word length, as well as the interaction between them, were treated as fixed factors. In addition, log word frequency was treated as a fixed factor in order to account for any differences between the mean word frequencies in each sentence type. Participants and items were treated as random factors, and a full random structure was used, although random slopes were sometimes removed in cases when the full model failed to converge.

The means and standard deviations for these measures are displayed in Table B3, and the LME output in Table B4. There was a clear tendency across all measures for longer words to receive longer fixations. This effect was significant in both gaze durations and total times, and marginal in single fixation durations. While there were clear trends for words in the uniform sentences to receive longer fixations than those in the non-uniform sentences, this effect did not reach significance in any of the measures. However, there was a marginal interaction between sentence uniformity and word length in total times (see Figure B1). While words of all three lengths received longer total times in the uniform than non-uniform sentences, the difference was larger (at least numerically) for the five letter words (40 ms) than the three (21 ms) or four letter words (17 ms). The trend for individual words within the uniform sentences to receive longer fixations than words of the same length in non-uniform sentences is in line with the finding that the uniform sentences were generally read more slowly than non-uniform sentences.

While these effects could be seen to suggest that there was some sort of inhibitory effect of within-sentence word length uniformity on lexical processing, we believe there is a simple explanation for these findings. As mentioned in the main body of Chapter 4 our uniform sentences were rated as being significantly less natural sounding than our non-uniform sentences. We maintain that this confound could not have made a significant contribution to our observed pattern of effects for skipping probabilities and saccadic targeting. However, in the case of our reading measures it seems plausible that reduced naturalness would have led to the trend towards increased reading times for all three word lengths in uniform relative to non-uniform sentences.

Table B3

Fixation Landing Position, Local Reading Time Measures, and Skipping Probabilities as a Function of Word Length, Sentence Uniformity, and the Difference (D) Between each Word Length in Uniform and Non-uniform Sentences.

	Three letter words			Four letter words			Five letter words		
	U	NU	D	U	NU	D	U	NU	D
FFD	224(72)	217(69)	7	229(75)	222(76)	7	234(74)	220(67)	14
GD	238(89)	227(82)	11	247(95)	238(93)	9	259(98)	244(93)	15
SFD	226(72)	219(71)	7	231(77)	222(75)	9	238(74)	222(68)	16
TT	275(129)	254(120)	21	282(134)	265(128)	17	320(161)	280(137)	40

Note. U= uniform; NU= non-uniform; D= difference; FFD = first fixation duration; GD= gaze duration; TT= total time; SFD= single fixation duration.

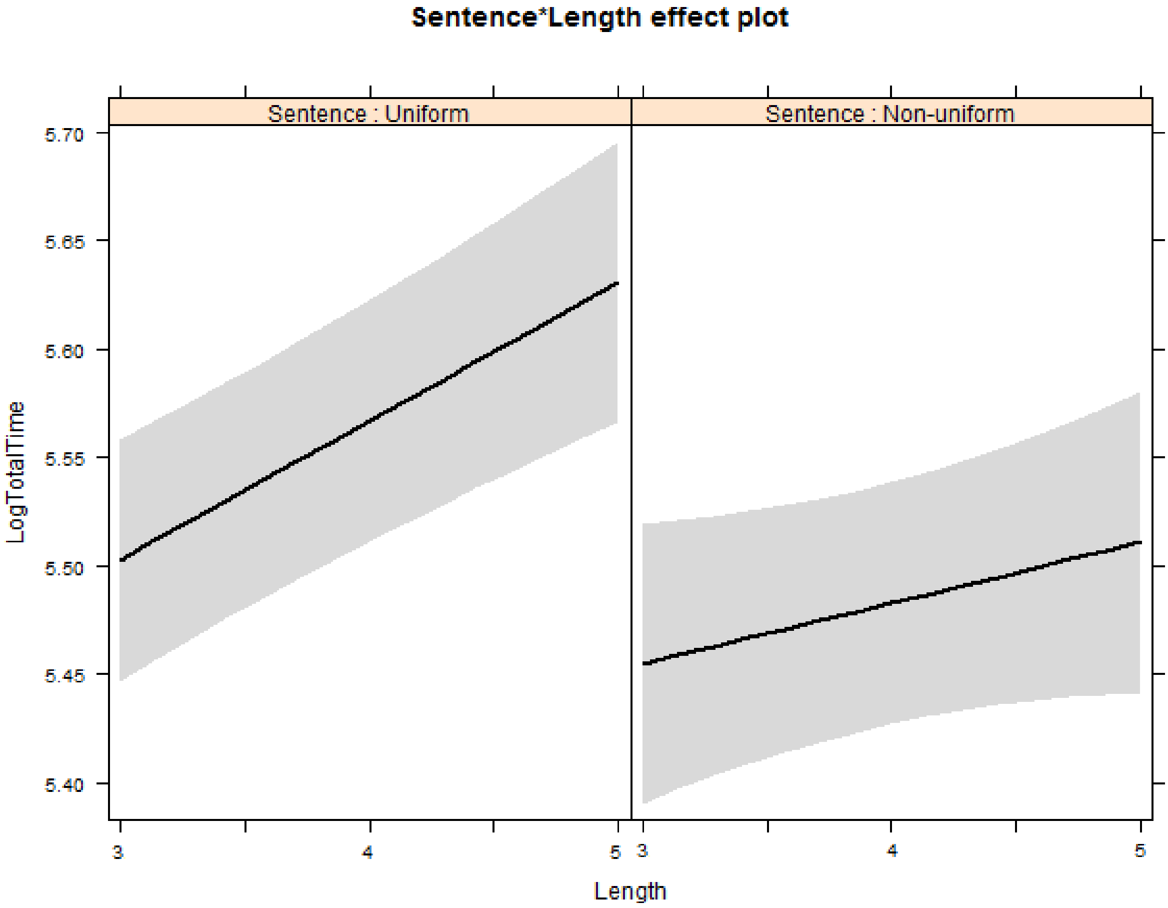


Figure B1. Fixed effects estimates for log transformed Total Time spent fixated on a word as a function of word length and sentence uniformity. 95% confidence bands are presented around the lines of estimated effects.

Table B4

Linear Mixed Model Analyses for Local Measures of Reading.

	FFD			GD			TT			SFD		
Model	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
Intercept	5.36	0.05	113.26	5.39	0.05	105.99	5.41	0.07	74.06	5.37	0.05	106.66
SU	0.03	0.05	0.66	0.01	0.07	0.14	0.06	0.08	0.76	0.05	0.06	0.88
WL	0.01	0.01	1.48	**0.03**	**0.01**	**3.36**	**0.05**	**0.01**	**5.33**	*0.01*	*0.01*	*1.83*
Frequency	-0.00	0.00	-1.48	**-0.01**	**0.00**	**-3.59**	**-0.01**	**0.00**	**-1.97**	**-0.00**	**0.00**	**-2.03**
SU*WL	-0.02	0.01	-1.33	-0.01	0.02	-0.80	*-0.04*	*0.02*	*-1.83*	-0.02	0.02	-1.60

Note. Significant terms are presented in bold. FFD = first fixation duration; GD= gaze duration; TT= total time; SFD= single fixation duration; SU= effect of moving from uniform to non-uniform sentences; WL= word length.

References

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish-English bilinguals. *Perception & Psychophysics, 63,* 875-890. doi:10.3758/BF03194444

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24,* 477-492. doi: 10.3758/BF03200936

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,* 1210-1220. doi:10.1037/a0023096

Angele, B., Schotter, E. R., Slattery, T. J., Tenenbaum, T. L., Bicknell, K., & Rayner, K. (2015). Do successor effects in reading reflect lexical parafoveal processing? Evidence from corpus-based and experimental eye movement data. *Journal of Memory and Language, 79–80,* 76-96. doi: 10.1016/j.jml.2014.11.003

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. doi:10.1080/13506280802009704

Angele, B., Tran, R., & Rayner, K. (2013). Parafoveal-on-foveal overlap can facilitate ongoing word identification during reading: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 526-538. doi: 10.1037/a0029492

Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements, *Language and Cognitive Processes, 19*, 391-426. doi: 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,* 416-424. doi: 10.1037/0278-7393.32.2.416

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

Balota, D. A., & Rayner, K. (1991). Word recognition processes in foveal and parafoveal vision: The range of influences of lexical variables. In D. Besner, & G.W.

Humphreys (Eds.). *Basic processes in reading* (pp. 198-232). Hillsdale, NJ: Erlbaum.

Balota, D. A., Yap, M. J., Hutchinson, K. A., Cortese, M. J., Kessler, B., Loftis, B., Nelly, J. H., Nelson, D. L., Simpson, G. D., & Treiman, R. (2007). The English lexicon project. *Behaviour Research Methods, 39*, 445-459. doi: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. doi: 10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes.* (R Package Version 0.999999–0). Retrieved from http://CRAN.R-project.org/package=lme4

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4 classes* (R Package Version 1.1-7). Retrieved from http://CRAN.R-project.org/package=lme4

Bélanger, N. N., Mayberry, R. I., & Rayner, K. (2013). Orthographic and phonological preview benefits: Parafoveal processing in skilled and less-skilled deaf readers. *Quarterly Journal of Experimental Psychology*, *66*, 2237-2252. doi:10.1080/17470218.2013.780085

Berniker, M., Voss, M., & Kording, K. (2010). Learning priors for Bayesian computations in the nervous system. *PLoS One*, *5*: e12686. doi:10.1371/journal.pone.0012686

Bertram, R., & Hyönä, J. (2007). The interplay between parafoveal preview and morphological processing in reading. In R. P. G. van Gompel, M. H. Fisher, W. S. Murray, & R. L. Hill (Eds.). *Eye movements: A window on mind and brain* (pp. 391-407). Oxford, UK: Elsevier.

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

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

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behaviour Research Methods, 41*, 977-990. doi:10.3758/BRM.41.4.977

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*, 711-733. Retrieved from: http://www.jstor.org/stable/4490266

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

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics, 32,* 45-61. doi: 10.1017/S0267190512000074

Cui, L., Yan, G., Bai, X., Hyönä, J., Wang, S., & Liversedge, S. P. (2013). Processing of compound-word characters in reading Chinese: An eye-movement-contingent display change study. *The Quarterly Journal of Experimental Psychology, 66*, 527-547. doi: 10.1080/17470218.2012.667423

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced compounds. *Journal of Experimental Psychology: Learning Memory and Cognition, 40,* 1778-1786. doi:10.1037/xlm0000013

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2015). How is information integrated across fixations in reading? In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 245-260). New York, NY: Oxford University Press.

Dare, N., & Shillcock, R. (2013). Serial and parallel processing in reading: Investigating the effects of parafoveal orthographic information on nonisolated word recognition. *Quarterly Journal of Experimental Psychology, 66*, 487-504. doi: 10.1080/17470218.2012.703212

Davis, C. J., & Bowers, J. S. (2004). What do letter migration errors reveal about letter position coding in visual word recognition? *Journal of Experimental Psychology: Human Perception and Performance, 30*, 923-941. doi: 10.1037/0096-1523.30.5.923

Deutsch, A., Frost, R., Pelleg, S., Pollatsek, A., & Rayner, K. (2003). Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin & Review, 10,* 415-422. doi: 10.3758/BF03196500

Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language and Cognitive Processes, 15,* 487-506. doi: 10.1080/01690960050119670

Deutsch, A., Frost, R., Pollatsek, A., Rayner, K. (2005). Morphological preview benefit effects in reading: Evidence from Hebrew. *Language and Cognitive Processes, 20,* 341-371. doi: 10.1080/01690960444000115

Drieghe, D. (2011). Parafoveal-on-foveal effects on eye movements during reading. In S. Liversedge, I. Gilchrist, & S. Everling (eds.) *Oxford Handbook on Eye Movements* (pp. 839-855). Oxford, UK: Oxford University Press.

Drieghe, D., Pollatsek, A., Juhasz, B. J., & Rayner, K. (2010). Parafoveal processing during reading is reduced across a morphological boundary. *Cognition, 116,* 136-142. doi:10.1016/j.cognition.2010.03.016

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

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

Engbert, R., & Krügel, A. (2010). Readers use Bayesian estimation for eye movement control. *Psychological Science, 21,* 366-371. doi:10.1177/0956797610362060

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*, 621–636. doi: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,* 777-813. doi: 10.1037/0033-295X.112.4.777

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text and Talk, 20,* 29-62. doi: 10.1515/text.1.2000.20.1.29

Fitzsimmons, G., & Drieghe, D. (2011). The influence of number of syllables on word skipping during reading, *Psychonomic Bulletin & Review, 18,* 736-741. doi: 10.3758/s13423-011-0105-x

Fitzsimmons, G., & Drieghe, D. (2013). How fast can predictability influence word skipping during reading? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1054–1063. doi:10.1037/a0030909

Fox, J., Weisberg, S., Friendly, M., & Hong, J. (2015). *Effect Display for Linear, Generalized Linear, and Other Models* (R Package Version, 3.0-4). Retrieved from http://CRAN.R-project.org/package=effects

Frost, R. (2015). Cross-linguistic perspective on letter-order processing: Empirical findings and theoretical considerations. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 88-98). New York, NY: Oxford University Press.

Grainger, J., Dufau, S., & Ziegler, J. C. (2016). A vision of reading. *Trends in Cognitive Sciences.* Advance online publication. doi: 10.1016/j.tics.2015.12.008

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. doi: 10.1016/j.actpsy.2013.11.014

Gu, J., & Li, X. (2015). The effect of character transposition within and across words in Chinese reading. *Attention, Perception, & Psychophysics, 77,* 272-281. doi: 10.3758/s13414-014-0749-5

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

Hautala, J., Hyönä, J., & Aro, M. (2011). Dissociating spatial and letter-based word length effects observed in readers' eye movement patterns. *Vision Research, 51,* 1719-1727. doi: 10.1016/j.visres.2011.05.015

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,* 417-429. doi: 10.1037/0278-7393.16.3.417

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. doi:10.1037/h0078814

Hintzman, D. L. (1984). MINERA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers, 16,* 96-101.

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 166-190. doi: 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,* 1150-1170. doi: 10.1037/a0020233

Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese.* Hillsdale, NJ: Erlbaum.

Hyönä, J., & Häikiö, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis. *Scandinavian Journal of Psychology, 46,* 475-483. doi: 10.1111/j.1467-9450.2005.00479.x

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eyemovement-contingent display change study. *Memory & Cognition, 32,* 523-532. doi:10.3758/BF03195844

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

Inhoff, A. W. (1989b). Lexical access during eye fixations in reading: Are word access codes used to integrate lexical information across interword fixations? *Journal of Memory and Language, 28,* 441-461. doi:10.1016/0749-596X(89)90021-1

Inhoff, A. W. (1990). Integrating information across eye fixations in reading: The role of letters and word units. *Acta Psychologia, 73,* 281-297. doi:10.1016/0001-6918(90)90027-D

Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuo-spatial attention and saccade programming during reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.). *Reading as a perceptual process* (pp. 221-246). Amsterdam, Netherlands: North-Holland/Elseiver Science Publisher. doi: 10.1016/B978-008043642-5/50012-7

Inhoff, A. W., Starr, M., & Shidler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception & Psychophysics, 62,* 1474-1484. doi: 10.3758/BF03212147

Johnson, R. L. (2007). The flexibility of letter coding: Nonadjacent letter transposition effects in the parafovea. In R. P. G. van Gompel, M. H. Fisher, W. S. Murray, & R. L. Hill (Eds.). *Eye movements: A window on mind and brain* (pp. 425-440). Oxford, United Kingdom: Elsevier.

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

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,* 209-229. doi: 10.1037/0096-1523.33.1.209

Juhasz, B. J., Pollatsek, A., Hyönä, J., Drieghe, D., & Rayner, K. (2009). Parafoveal processing within and between words. *The Quarterly Journal of Experimental Psychology, 62,* 1356-1376. doi:10.1080/17470210802400010

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

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

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

Just, M. A., & Carpenter, P. A. (1990). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329-354. doi: 10.1037/0033-295X.87.4.329

Kaakinen, J. K., & Hyönä, J. (2014). Task relevance induces momentary changes in the functional visual field during reading. *Psychological Science, 25*, 626-632. doi: 10.1177/0956797613512332

Kambe, G. (2004). Parafoveal processing of prefixed words during eye fixations in reading: Evidence against morphological influences on parafoveal preprocessing. *Perception & Psychophysics, 66,* 279-292. doi: 10.3758/BF03194879

Kapoula, Z. (1985). Evidence for a range effect in the saccadic system. *Vision Research, 25,* 1155-1157. doi: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,* 735-743. doi: 10.1016/0042-6989(86)90087-8

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

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,* 262-284. doi: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,* 12–35. doi: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,* 1250-1255. doi:10.1037/0096-1523.33.5.1250

Krügel, A., & Engbert, R. (2014). A model of saccadic landing positions in reading under the influence of sensory noise. *Visual Cognition, 22,* 334-353. doi:10.1080/13506285.2014.894166

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

Lima, S. D. (1987). Morphological analysis in sentence reading. *Journal of Memory and Language, 26,* 84-99. doi:10.1016/0749-596X(87)90064-7

Liu, W., Inhoff, A. W., Ye, Y., & Wu, C. (2002). Use of parafoveally visible characters during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1213-1227.

Liversedge, S. P., Drieghe, D., Li, X., Yan, G., Bai, X., & Hyönä, J. (2016). Universality in eye movements and reading: A trilingual investigation. *Cognition, 147,* 1-20. doi: 10.1016/j.cognition.2015.10.013

Liversedge, S. P., & Findlay, J.M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science, 4,* 6-14. doi: 10.1016/S1364-6613(99)01418-7

Liversedge, S. P., Rayner, K., White, S. J., Vergilion-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*, 1013-1024. doi: 10.1016/j.visres.2003.12.002

Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition, 126*, 313-318. doi: 10.1016/j.cognition.2012.09.010

Martin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81,* 899-917.

McClelland, J. L., & Rumelheart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88,* 375-407. doi: 10.1037/0033-295X.88.5.375

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

McConkie, G. W., & Rayner, K. (1976a). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society, 8,* 365-368. doi: 10.3758/BF03335168

McConkie, G. W., & Rayner, K. (1976b). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer & R. B. Ruddell (Eds.). *Theoretical models and processes of reading* (pp. 137-162). Newark, NJ: International Reading Association.

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

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, 1107-1118. doi: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, 245-253. doi:10.3758/BF03208086

McDonald, S. A. (2006). Effects of number-of-letters on eye movements during reading are independent from effects of spatial word length. *Visual Cognition, 13,* 89-98. doi: 10.1080/13506280500143367

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*, 721–728. doi:10.1111/j.1467-9280.2009.02364.x

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

Mozer, M. C. (1983). Letter migration in word perception. *Journal of Experimental Psychology: Human perception and performance*, *9*, 531. doi: 10.1037/0096-1523.9.4.531

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

O'Regan, J. K., & Jacobs, A. M. (1992). Optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance, 18.* 185-197. doi:10.1037/0096-1523.18.1.185

O'Regan, J. K., Lévy-Schoen, A., Pynte, J., & Brugaillère, B. É. (1984). Convenient fixation location within isolated words of different length and structure. Journal of Experimental Psychology: Human Perception and Performance, 10, 250-257. doi:10.1037/0096-1523.10.2.250

Pagán, A., Blythe, H. I., & Liversedge, S. P. (2015). Parafoveal preprocessing of word initial trigrams during reading in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication. doi: 10.1037/xlm0000175

Paterson, K. B., Almabruk, A. 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.* Advance online publication. doi:10.3758/s13423-015-0809-4

Perea, M., & Carreiras, M. (2012). Perceptual uncertainty is a property of the cognitive system. *Behavioral and Brain Sciences*, *35*, 298-299. doi:10.1017/S0140525X12000118

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, *108,* 3526-3529. doi: 10.1073/pnas.1012551108

Pinker, S. (1999). *Words and rules: The ingredients of language*. New York, NY: Harper-Collins.

Plummer, P., & Rayner, K. (2012). Effects of parafoveal word length and orthographic features. *Attention, Perception, & Psychophysics, 74*, 950–963. doi:10.3758/s13414-012-0286-z

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The role of morphological constituents in reading Finnish compound words. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 820-833. doi: t0.1037//0096-1523.26.2.820

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,* 148-162. doi: 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*, 1142–1158. doi:10.1037/0096-1523.25.4.1142

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*, 1–56. doi:10.1016/j.cogpsych.2005.06.001

R Development Core Team. (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, Retrieved from http://www.Rproject.org/

Radach, R., Inhoff, A. W., Glover L., & Vorstius, C. (2013). Contextual constraint and n+2 preview effects in reading. *The Quarterly Journal of Experimental Psychology,* doi:10.1080/17470218.2012.761256

Radach, R., Inhoff, A., & Heller, D. (2004). Orthographic regularity gradually modulates saccade amplitudes in reading. *European Journal of Cognitive Psychology, 16*, 27–51. doi:10.1080/09541440340000222

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

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

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

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. The quarterly journal of experimental psychology, 62(8), 1457-1506. doi:10.1080/17470210902816461

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

Rayner, K. & Bertera, J. H. (1979). Reading without a fovea. *Science, 206*, 468-469. doi: 10.1126/science.504987

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

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*, 230-245. doi:10.1037/0096-1523.33.1.230

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*, 310-323. doi: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,* 385-388. doi:10.1111/1467-9280.24483

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

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

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*, 1188-1200. doi: 10.1037/0096-1523.22.5.1188

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

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

Reichle, E. D., & Drieghe, D. (2014). Using E-Z Reader to examine the consequences of fixation-location measurement error. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 262-270. doi: 10.1037/a0037090

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

Reichle, E. D., & Perfetti, C. A. (2003). Morphology in word identification: A word-experience model that accounts for morpheme frequency effects. *Scientific Studies of Reading, 7*, 219-237. doi: 10.1207/S1532799XSSR0703_2

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

Reichle, E. D., Pollatsek, A., Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, *7,* 4-22. doi:10.1016/j.cogsys.2005.07.002

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: accounting for initial fixation locations and refixation within the E-Z Reader model. *Vision Research, 39,* 4403-4411. doi:10.1016/S0042-6989(99)00152-2

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26,* 445-476. doi:10.1017/S0140525X03000104

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

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology, 65,* 177-206. doi:10.1016/j.cogpsych.2012.03.001

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,* 1026-1042. doi: 10.1037/a0027735

Risse, S., & Kliegl, R. (2014). Dissociating preview validity and preview difficulty in parafoveal processing of word *n*+1 during reading. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 653-668. doi:10.1037/a0034997

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

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

Schotter, E R., Bicknell, K., Howard, I., Levy, R., & Rayner, K. (2014). Task effects reveal cognitive flexibility responding to frequency and predictability: Evidence from eye movements in reading and proofreading. *Cognition, 131*, 1-27. doi:10.1016/j.cognition.2013.11.018

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

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, and Psychophysics, 74,* 5-35. doi: 58/s13414-011-0219-2

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, 309-333. doi: 10.1080/13506285.2013.873508

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 776-784. doi:10.1037/a0022531

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,* 1924-1938. doi:10.1037/a0024322

Treisman, A., & Souther, J. (1986). Illusory words: The roles of attention and of top-down constraints in conjoining letters to form words. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 3-17. doi:10.1037/0096-1523.12.1.3

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

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. doi:10.1037/0096-1523.34.1.205

White, S. J., & Liversedge, S. P. (2006a). Foveal processing difficulty does not modulate non-foveal orthographic influences on fixation positions. *Vision Research, 46,* 426-437. doi:10.1016/j.visres.2005.07.006

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

White, S. J., Bertram, R., & Hyönä, J. (2008). Semantic processing of previews within compound words. *Journal of Experimental Psychology: Learning Memory and Cognition, 34,* 988-993. doi: 10.1037/0278-7393.34.4.988

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

Williams, R., & Morris, R. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology, 16,* 312–339. doi:10.1080/09541440340000196

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*, 1669-1676. doi: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,* 561-566. doi: :10.3758/PBR.16.3.561

Yan, M., Sommer, W. (2015). Parafoveal-on-foveal effects of emotional word semantics in reading Chinese sentences: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1237-1243. doi:10.1037/xlm0000095

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,* 1069-1075. doi:10.1037/a0026935

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,* 181-215. doi:10.1016/j.cognition.2014.03.008

Yao-N'Dré, M., Castet, E., & Vitu, F. (2014). Inter-word eye behaviour during reading is not invariant to character size: Evidence against systematic saccadic range error in reading. *Visual Cognition, 22,* 415-440. doi: 10.1080/13506285.2014.886652

Yu, L., Cutter, M. G., Yan, G., Bai, X., Fu, Y., Drieghe, D., & Liversedge, S. P. (2016). Word n+2 preview effects in three-character Chinese idioms and phrases. *Language, Cognition and Neuroscience*. Advance Online Publication. doi: 10.1080/23273798.2016.1197954

Zhou, X., & Marslen-Wilson, W. (1999). Sublexical processing in reading Chinese. In J. Wang, A. W. Inoff, & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 37-64). Mahwah, NJ: Erlbaum.