COPY OF:

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.

Running head: WORD FREQUENCY AND ORTHOGRAPHY

Eye movement control during reading:

Effects of word frequency and orthographic familiarity.

Sarah J. White, University of Leicester

School of Psychology

University of Leicester

Lancaster Road

Leicester

LE1 9HN

UK

Tel: +44 (0)116 229 7181

Fax: +44 (0)116 229 7196

E-mail: s.j.white@dunelm.org.uk

# Abstract

Word frequency and orthographic familiarity were independently manipulated as readers' eye movements were recorded. Word frequency influenced fixation durations and the probability of word skipping when orthographic familiarity was controlled. These results indicate that lexical processing of words can influence saccade programming (as shown by fixation durations and which words are fixated). Orthographic familiarity, but not word frequency, influenced the duration of prior fixations. These results provide evidence for orthographic, but not lexical, parafoveal-on-foveal effects. Overall, the findings have a crucial implication for models of eye movement control in reading: there must be sufficient time for lexical factors to influence saccade programming before saccade metrics and timing are finalised. The conclusions are critical for the fundamental architecture of models of eye movement control in reading, namely, how to reconcile long saccade programming times and complex linguistic influences on saccades during reading.

Key words / phrases: Reading. Eye movements. Word frequency. Orthography. Models of eye movement control in reading.

As we read, the linguistic characteristics of words influence the duration of fixations and which words are fixated (Rayner, 1998). The present study provides a detailed examination of the influences of orthographic familiarity and word frequency on eye movements during reading. It thus provides a critical assessment of the relationship between linguistic text processing and the systems that control when and where the eyes move. Specifically, whether lexical processing can have an immediate influence on saccade programming is examined. The issues addressed here have crucial implications for the architecture of models of eye movement control in reading. The article adds to a growing number of recent studies which specifically aim to test and develop such accounts (Inhoff, Eiter, & Radach, 2005; Kliegl, Nuthmann, & Engbert, 2006; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner, Juhasz, & Brown, 2007; Rayner, Liversedge, White, & Vergilino-Perez, 2003; Reingold & Rayner, 2006). Previous studies of the effects of word frequency and orthography will first be discussed. Models of eye movements in reading will be summarised which may account for such effects. Finally, the issue of whether processing of parafoveal information can influence prior fixations (parafoveal-on-foveal effects) and where words are fixated (saccade specification) will be considered.

*Word frequency effects*

The influence of word frequency on word processing is an established finding both for isolated word response time tasks (Monsell, 1991) and sentence reading (Rayner, 1998). Inhoff and Rayner (1986) (also Rayner & Duffy, 1986) first demonstrated, that for words in sentences with word length controlled, first fixation durations and gaze durations (the sum of fixations before leaving a word) are longer on infrequent than frequent words. Word frequency effects during sentence reading are usually spatially localised to the word

that induced those effects (Henderson & Ferreira, 1990, 1993; Raney & Rayner, 1995) and in other cases there are effects of word frequency both on the word itself and on subsequent spillover fixations (Kennison & Clifton, 1995; Kliegl et al., 2006; Rayner & Duffy, 1986). Frequent words are also more likely to be skipped than infrequent words (e.g. Rayner, Sereno, & Raney, 1996; see Brysbaert, Drieghe, & Vitu, 2005; Brysbaert & Vitu, 1998). Note that other factors, such as age of acquisition and concreteness, have been shown to have independent influences on reading behaviour even though they are correlated with word frequency (Juhasz & Rayner, 2003).

The influence of word frequency on eye movement behaviour during reading suggests that lexical word recognition processes can influence when and where the eyes move. However, of the large number of studies that manipulated word frequency, a very small proportion of these have attempted to control for the orthographic characteristics of the words. It is easy to confound word frequency with orthographic familiarity because frequent words are necessarily orthographically familiar whereas many infrequent words are orthographically unfamiliar. A few studies have attempted to control for orthographic characteristics using the measure of type frequency, which is the number of words that contain a particular letter sequence. Rayner and Duffy (1986) undertook post-hoc analyses which showed higher trigram type frequencies for frequent than infrequent target words, but they found that word frequency effects still held for those items in which differences in type frequency were reversed. Bertram and Hyönä (2003) showed word frequency effects on first fixation and gaze durations when bigram type frequency was controlled. Rayner et al. (1996) reported effects of word frequency on fixation durations and word skipping for items which had equal monogram and bigram type frequency counts.

Importantly, previous reading studies only attempted to control for orthography by using type frequency counts. As type frequency is simply the number of words that contain a particular letter sequence, it is effectively a measure of lexical informativeness or redundancy. For example, the trigram "*pne*" at the word beginning has a very low type frequency and is very informative because it highly constrains the number of possible word candidates (*pneumatic*, *pneumonia*). Importantly, type frequency does not reflect letter sequence familiarity. For example, there are a number of words that begin with the letter sequence "irr", but very few of these are very frequent; hence although "irr" has a relatively high type frequency it actually has quite low orthographic familiarity. Manipulations of only type frequency or informativeness should therefore reflect processing at the level of lexical candidates.

A better measure of orthographic familiarity is token frequency, which is the sum of the frequencies of words that contain a particular letter sequence[1]. Critically, effects of orthographic familiarity may involve processing at a sub-lexical or even visual level (see next section). Note that other studies have tested for effects of word frequency, and the type and token frequency of word initial letters, in multiple isolated word processing tasks (Kennedy, 1998; 2000; Kennedy, Pynte, & Ducrot, 2002) and sentence reading (Kennedy & Pynte, 2005; Pynte & Kennedy, 2006). However the orthographic characteristics beyond the word initial letters were not controlled in these studies and consequently they do not test whether word frequency effects occur independently of differences in orthographic familiarity. To summarise, although previous studies of word frequency have attempted to control for orthography to some extent, these studies do not eliminate the

possibility that differences in orthographic familiarity could have produced, or at least influenced, the word frequency effect.

The issue of whether orthographic familiarity may be contributing to the word frequency effect is of particular importance for the case of word skipping. Critically, the linguistic characteristics of skipped words must be processed in parafoveal vision (where stimuli are degraded due to acuity limitations) which reduces the speed of linguistic processing of those words (Rayner & Morrison, 1981; Schiepers, 1980). In addition, fixations prior to word skipping are likely to involve processing of the fixated word too; which may reduce the time (or resources) available within a fixation for processing of the parafoveal word (e.g. Morrison, 1984). Of the studies that have shown word frequency effects on word skipping, only one study controlled for type bigram frequency (Rayner et al., 1996). Not only was orthographic familiarity not controlled, but orthographic processing of trigrams, not just bigrams, could have contributed to the effect. Furthermore, the fact that some studies have not shown word frequency effects on word skipping (Calvo & Meseguer, 2002; Henderson & Ferreira, 1993) raises the possibility that other factors, such as orthographic familiarity, may be important. Visual familiarity may play an important role in word processing (Martin, 2004) and some models suggest that fixation durations (McDonald, Carpenter, & Shillcock, 2005) and especially word skipping (Engbert, Nuthmann, Richter, & Kliegl, 2005) are not necessarily driven by full word identification. Consequently it is particularly feasible that processing of orthographic familiarity may have produced the previously reported effects of word frequency on word skipping.

To summarise, previous studies have not provided a proper test of whether word frequency influences fixation durations and word skipping when orthographic familiarity is fully controlled. Differences in orthographic familiarity may have caused, or at least inflated, effects that have been attributed to word frequency. Consequently previous studies of word frequency do not categorically demonstrate lexical influences on word recognition and eye movement behaviour. The issue of whether word frequency effects may be explained by differences in orthographic processing is crucial. If the effects are due to differences in orthography then this raises the critical question as to whether lexical factors can have an immediate impact on when and where the eyes move during reading. Basically, given that this issue is so critical for models of eye movement control during reading, it is absolutely essential that a thorough analysis of the effects of word frequency and orthographic familiarity be undertaken. The present study does this by testing the effects of word frequency on eye movement behaviour whilst controlling for monogram, bigram and trigram token orthographic frequencies.

*Orthographic effects*

Studies using isolated word tasks have investigated a wide range of factors related to orthographic processing of words (Henderson, 1982). A number of early studies were undertaken into the effects of bigram frequency and word frequency during reading (Gernsbacher, 1984). These studies used isolated word methods such as tachistoscopic presentation of words, naming and lexical decision and bigram frequency was controlled only by type frequency counts such as those reported by Mayzner and Tresselt (1965). However the findings of such studies using low frequency words have been contradictory with some indicating that words with high frequency bigrams are more difficult to process

than those with low frequency bigrams (Broadbent & Gregory, 1968; Rice & Robinson, 1975) whilst others show the opposite result (Biederman, 1966; Seidenberg, Waters, Barnes, & Tanenhaus, 1984; Waters & Seidenberg, 1985). However as noted above, letter sequences can have different levels of informativeness and familiarity and it is possible that such differences may be one explanation for the inconsistent results (Grainger & Dijkstra, 1996).

A number of studies using eye tracking methodologies have examined the role of the informativeness of letter sequences (confounded with orthographic familiarity). Pynte, Kennedy and Murray (1991) showed longer fixation durations on informative parts of words (see also Holmes & O'Regan, 1987). In sentence reading, Lima and Inhoff (1985) showed that first fixation durations were longer on words which began with constraining (e.g. *dwarf*) compared to less constraining (e.g. *clown*) letter sequences. Other experiments have shown that the orthographic familiarity of the initial letters of long words can influence where they are first fixated (Hyönä, 1995; Radach, Inhoff, & Heller, 2004; Vonk, Radach, & van Rijn, 2000; White & Liversedge, 2004, 2006a, 2006b).

Orthographic familiarity may impact on visual, sub-lexical or lexical levels of processing. Processing of text at a visual, rather than linguistic, level may modulate the familiarity of visual information such that frequent letter strings may develop higher visual familiarity than infrequent letter strings (Findlay & Walker, 1999). Therefore any effects of orthographic familiarity must be interpreted as a reflection of processing at least at the level of visual familiarity. However differences in orthographic familiarity may also be associated with differences in informativeness or constraint (type frequency) and orthographically unfamiliar words may also tend to have more irregular phonology and

fewer orthographic neighbours. Consequently, any effects of orthographic familiarity

could also be driven by sub-lexical or even lexical processes. The present study

manipulates the orthographic familiarity of the entire word, and it examines whether

orthographic familiarity influences fixation durations and word skipping.

*Models of eye movement control during reading*

Models of eye movement control vary in the extent to which they suggest that

linguistic processing can influence eye movement behaviour. Some suggest that eye

movements are driven by linguistic processing at least at the level of lexical access (Just &

Carpenter, 1980; Morrison, 1984; Thibadeau, Just, & Carpenter, 1982). Others suggest

that linguistic processing influences, or even determines, when and where the eyes move,

but this may be indexed by an early stage of word processing (Engbert, Longtin, & Kliegl,

2002; Engbert et al., 2005; Feng, 2006; Kliegl & Engbert, 2003; Legge, Hooven, Klitz,

Mansfield, & Tjan, 2002; Legge, Klitz, & Tjan, 1997; McDonald et al., 2005; Pollatsek,

Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek,

& Rayner, 2006; Reichle, Rayner, & Pollatsek, 1999, 2003; Reilly & Radach, 2003, 2006;

Richter, Engbert, & Kliegl, 2006).

In contrast, it has been suggested (Deubel, O'Regan, & Radach, 2000; Nazir,

2000), and models have proposed (McConkie & Yang, 2003; O'Regan, 1990, 1992; Reilly

& O'Regan, 1998; Suppes, 1990; Yang & McConkie, 2001; Yang, 2006), that eye

movements are largely controlled by visual and oculomotor factors and that linguistic

processes have a smaller influence, for example by delaying or cancelling programmed

saccades. However such models can not necessarily account for some linguistic influences

on reading behaviour, such as the following frequency effects: on the first of multiple first

pass fixations on a word (Rayner et al., 1996); in the absence of visual information (Liversedge, Rayner, White, Vergilino-Perez, Findlay, & Kentridge, 2004; Rayner, Liversedge et al., 2003); or on gaze durations that can not be explained by a number of very long fixations or gaze durations (Rayner, 1995; Rayner, Liversedge et al., 2003). Linguistic factors therefore have a critical role in influencing eye movement behaviour during reading and visual and/or oculomotor based models can not fully account for reading eye movement behaviour. However, the issue of precisely how linguistic processing influences when and where the eyes move during reading is far from resolved.

The most comprehensive implemented models of eye movement behaviour during reading, the E-Z reader model (Pollatsek et al., 2006; Reichle et al., 1998, 1999, 2003, 2006) and the SWIFT model (Engbert et al., 2002, 2005; Kliegl & Engbert, 2003; Richter et al., 2006), include a period of saccade programming during which the planned timing and metrics of the subsequent saccade are prepared (see also Salvucci, 2001). The time necessary to programme a saccade has been estimated to be 175-200ms (Rayner, Slowiaczek, Clifton, & Bertera, 1983), or 125ms once 50 ms of mandatory visual processing has been taken into account (Pollatsek et al., 2006). Fixation durations in reading are on average 200-250ms long (Rayner, 1998) and so a substantial portion of the time during a fixation could involve programming the next saccade. Although linguistic processing can continue during this period, it may not necessarily influence saccade programming.

Saccade programming is composed of a labile period, during which the saccade can be cancelled, and a non-labile period, during which the saccade metrics are finalized and the saccade can not be cancelled (Becker & Jürgens, 1979; Deubel, O'Regan, &

Radach, 2000). In the E-Z reader model the mean durations of the labile and non-labile

stages are 100 and 25ms respectively (Pollatsek et al., 2006). In the SWIFT model the

labile stage is between 50 and 150ms and the non-labile stage is between 5 and 50ms

(Engbert et al., 2005) or else these values are fixed at 150 and 50ms respectively (Richter

et al., 2006). In the E-Z reader and SWIFT models, linguistic factors can determine, or

influence, when the labile stage of saccade programming commences. In both cases, the

time at which the saccade programme is triggered effectively influences when the

subsequent saccade is executed, and therefore the duration of fixations. Critically, given

the time required for saccade programming (~125ms) and the average duration of fixations

(~250ms) such linguistic influences on when saccade programming is initiated would

necessarily have to occur relatively early during fixations.

Linguistic processing during the labile stage of saccade programming may also

influence current fixation durations or the next saccade target. In both the E-Z reader and

SWIFT models, such mechanisms have been adopted to account for word skipping. In the

E-Z reader model, the labile stage can be cancelled and re-started, such that the word

target is changed and a new saccade programme can commence. In contrast, in the SWIFT

model the saccade target location is always specified at the end of the labile stage. Note

that even when linguistic factors impact late during a fixation at the end of the labile stage

of saccade programming, there would still be some delay before the saccade could be

executed due to the subsequent non-labile stage.

The issue of how linguistic processing can influence saccade programming, whilst

taking account of saccade programming times, is one of the most critical issues for the

design of models of eye movements in reading. The present study tests whether lexical

factors (word frequency) influence saccade programming (fixation durations and word skipping) independent of sub-lexical factors (orthographic familiarity). If the lexical factor of word frequency influences fixation durations or word skipping, then this must either be explained by an early linguistic influence on when saccade programming is initiated, or else a later influence during, or at the end of, the labile stage of saccade programming.

Both the E-Z reader and SWIFT models suggest that word frequency impacts on saccade programming, and so they predict that this influences both fixation durations and word skipping. However, both models are vague about precisely how word frequency impacts on the difficulty of word processing. Given the timing constraints of the proposal that linguistic factors trigger the initiation of saccade programming, it is quite plausible that the sub-lexical factor of orthographic familiarity, and not the lexical factor of word frequency, might influence saccades within such a framework. Nevertheless, if the present study shows that word frequency does have an influence, for example on how long words are first fixated, then this would suggest that the linguistic processes that occur prior to saccade programming in these models must be of a lexical nature.

*Parafoveal-on-foveal effects of word frequency and orthography*

Many studies have shown that parafoveal processing of words to the right of fixation can influence the probability of word skipping and can also facilitate subsequent processing of those words when they are later fixated, known as preview benefit (Rayner & Pollatsek, 1981). However, the issue of whether parafoveal processing of words can influence fixations prior to fixating those words (fixation n-1), known as "parafoveal-on-foveal effects", is controversial. Previous research testing whether orthographic familiarity and word frequency produce parafoveal-on-foveal effects is reviewed below. The

investigation of parafoveal-on-foveal effects is especially critical because it has important

implications for models of eye movements in reading. Serial processing models suggest

that words are lexically processed one at a time (Morrison, 1984; Pollatsek et al., 2006;

Reichle et al., 1998, 1999, 2003, 2006). In contrast, parallel accounts suggest that multiple

words can be processed at once (Engbert et al., 2002, 2005; Kennedy, 2000; Kliegl &

Engbert, 2003; McDonald et al., 2005; Reilly & Radach, 2003, 2006; Richter et al., 2006).

If the characteristics of a parafoveal word influence fixation n-1, this could be indicative

of parallel processing of words. There are alternative explanations for these effects, such

as inaccurate saccade targeting, which are explained below and further in the Discussion.

To be clear, parallel based models predict both sub-lexical and lexical parafoveal-on-

foveal effects, whereas the core assumptions of serial based models do not predict such

effects, although additional assumptions or explanations have been offered to account for

these. Importantly, the present study provides an important opportunity to test for the

presence of parafoveal-on-foveal effects in a carefully controlled experiment.

Some sentence reading studies suggest that parafoveal preprocessing at least at the

level of orthographic familiarity can influence fixation n-1 (Inhoff, Starr, & Shindler,

2000; Pynte, Kennedy, & Ducrot, 2004; Rayner, 1975; Starr & Inhoff, 2004; Underwood,

Binns, & Walker, 2000) though other studies have shown no such effects (Rayner, Juhasz

et al., 2007; White & Liversedge, 2004, 2006b). In addition, a number of studies using

multiple isolated word processing tasks (Kennedy, 1998, 2000; Kennedy, Pynte, &

Ducrot, 2002) and sentence reading (Kennedy & Pynte, 2005; Pynte & Kennedy, 2006)

have suggested that the informativeness of the word initial letters of a parafoveal word can

influence fixation n-1. Some studies using isolated word processing (Kennedy, 1998,

2000; Kennedy, Pynte, & Ducrot, 2002) and sentence reading (Kennedy & Pynte, 2005; Kliegl et al., 2006; Pynte & Kennedy, 2006) tasks have suggested that word frequency can also produce such effects. However other studies have shown inconsistent (Hyönä & Bertram, 2004) or no (Calvo & Meseguer, 2002; Henderson & Ferreira, 1993; Rayner, Fischer, & Pollatsek, 1998; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999) parafoveal-on-foveal effects of word frequency, and other research has shown no evidence of other lexical parafoveal-on-foveal effects (Altarriba, Kambe, Pollatsek, & Rayner, 2001; Hyönä & Häikiö, 2005; Inhoff et al., 2000; Rayner, 1975; Rayner, Juhasz et al., 2007) (for reviews see: Rayner, White, Kambe, Miller, & Liversedge, 2003; Rayner & Juhasz, 2004).

Many fixations are mislocated during reading (McConkie, Kerr, Reddix, & Zola, 1988; Nuthmann et al., 2005) and it has been suggested that parafoveal-on-foveal effects may arise due to such inaccurately targeted saccades (Drieghe, Rayner, & Pollatsek, 2007; Rayner, Warren et al., 2004; Rayner, White et al., 2003). That is, saccades intended to land on the critical word mistakenly land at the end of the previous word, but attention is still allocated to the originally intended location, such that processing of the critical word influences the fixation duration on the previous word. Importantly, these effects would be expected to be in a conventional direction. For example, lower frequency and less familiar parafoveal words would produce longer fixations on foveal words than higher frequency and more familiar parafoveal words. Reports of post-lexical parafoveal-on-foveal effects have shown effects of a conventional direction which might therefore be explained by inaccurate saccade targeting (Inhoff, Radach, Starr, & Greenberg, 2000; Murray, 1998; Rayner, Warren et al., 2004). Nevertheless, Starr and Inhoff (2004) showed that parafoveal-on-foveal effects of orthographic familiarity held even when prior fixations at

the very end of the foveal word were eliminated from analysis. Furthermore, other studies that have shown that parafoveal information influences the probability of refixating on the foveal word, or that the parafoveal-on-foveal effects are opposite to the conventional direction, can not be explained by the inaccurate saccade targeting account (Kennedy, 1998, 2000; Kennedy & Pynte, 2005; Kennedy, Pynte, & Ducrot, 2002).

The present study provides an additional test of whether sub-lexical (orthographic familiarity) and lexical (word frequency) characteristics of parafoveal words influence foveal fixations in the form of parafoveal-on-foveal effects. The question of whether there are lexical parafoveal-on-foveal effects in sentence reading under carefully controlled experimental conditions is particularly important because so far the only sentence reading studies to have shown parafoveal-on-foveal effects of word frequency have been based on corpus of reading data (Kennedy & Pynte, 2005; Pynte & Kennedy, 2006) or data from a combination of different studies (Kliegl et al., 2006). It is important to assess whether lexical parafoveal-on-foveal effects arise under standard experimental conditions when the stimuli are designed to control for other variables. In the present study sentence beginnings were identical across the conditions and none of the critical words was predictable from the sentence context. In addition, both foveal and parafoveal words were short, therefore providing optimal conditions for parafoveal preview, and consequently for parafoveal-on-foveal effects to occur.

*Saccade specification and regressions*

The current study also provides an opportunity to test whether orthographic familiarity and word frequency influence saccade length into short words or where short words are first fixated. As noted above, previous research suggests that long words with

orthographically unfamiliar word beginnings are first fixated nearer to the beginning of the word than long words with orthographically familiar beginnings (Hyönä, 1995; Radach et al., 2004; Vonk, et al. 2000; White & Liversedge, 2004, 2006a, 2006b) whereas Rayner et al. (1996) showed no effect of word frequency on where words are first fixated. Importantly, previous studies have not tested whether orthographic familiarity can influence where short words are first fixated. Note that models of eye movement control during reading generally predict that the only parafoveal information used to influence saccade specification to word targets is word length, hence they would predict no effects of orthographic familiarity or word frequency on initial fixation positions (e.g. Engbert et al., 2005; Reichle et al., 1999). Reichle et al. (2003) suggest that low spatial frequency information (such as ascenders and descenders) might influence saccade programming, however they do not specify exactly what influence this information would have on where words are fixated (Liversedge & White, 2003).

Finally, the experiment presented here also tests whether word frequency and orthographic familiarity influences the probability of making regressions out of, or into, words. Some models of eye movement control do not attempt to account for interword regressions (e.g. McDonald et al., 2005; Pollatsek et al., 2006; Reichle et al., 1998, 1999, 2003, 2006) whereas others suggest that regressions are made to words with incomplete lexical processing (Engbert et al., 2002, 2005; Kliegl & Engbert, 2003; Richter et al., 2006) (see also: Reilly & Radach, 2003, 2006). The present study investigates the relative influences of sub-lexical (orthographic familiarity) and lexical (word frequency) influences on interword regressions.

*Summary*

The present study provides the first comprehensive test of the independent effects of word frequency and (whole word) orthographic familiarity on word recognition processes, as shown by eye movement behaviour during reading. The main focus of the study is to test whether lexical processes affect saccade programming, such that fixation durations and word skipping are influenced by the lexical characteristics of words. The study also provides an opportunity to test whether the orthographic or lexical characteristics of parafoveal words might be processed in parallel with the fixated word, or at least produce parafoveal-on-foveal effects. Measures of saccade specification, such as where words are initially fixated, are also reported, as well as analyses of whether orthographic familiarity or word frequency influences regressions out of, or into, words.

In order to investigate the independent effects of word frequency and orthographic familiarity an eye movement experiment was undertaken with three conditions. Word frequency effects were examined by comparing frequent (e.g. *town*) and infrequent (e.g. *cove*) words that were equally orthographically familiar. Orthographic familiarity effects were examined by comparing orthographically familiar (e.g. *cove*) and unfamiliar words (e.g. *quay*) that were equally infrequent. That is, the orthographically familiar low word frequency condition was used for both analyses. It was not possible to orthogonally manipulate the variables because words with high frequencies can not be orthographically unfamiliar.

Method

*Participants*. Thirty students at the University of Durham participated in the experiment. They all had normal or corrected to normal vision and were naïve regarding the purpose of the experiment.

*Materials and Design*. The critical words were (1) frequent and orthographically familiar, (2) infrequent and orthographically familiar, or (3) infrequent and orthographically unfamiliar; these three conditions were manipulated within participants and items.

Word frequencies and n-gram frequencies were calculated in counts per million using the CELEX English word form corpus (Baayen, Piepenbrock, & Gulikers, 1995). The word form, rather than lemma, corpus was selected in order to ensure inclusion of letter sequences within all words. For example, the word form corpus includes words such as "*went*", whereas in the lemma corpus frequencies for such words would be associated instead with the base form (e.g. "*go*"). For the orthographically familiar words, the frequent words had higher word frequency ($M = 297$, $SD = 166$) than the infrequent words ($M = 1.7$, $SD = 0.9$), $t(38) = 11.1$, $p < .001$. For the infrequent words, there was no difference in word frequency between the orthographically unfamiliar ($M = 1.5$, $SD = 1$) and familiar ($M = 1.7$) conditions ($t < 1$).

Previous research indicates that letter position coding is quite flexible (e.g. Humphreys, Evett, & Quinlan, 1990). Consequently n-gram frequencies were calculated both specific to position and non-position specific regardless of word length. Orthographic familiarity was measured using n-gram token frequencies, which represent the sum of the frequencies of the words that contain a particular letter sequence. Token frequencies for

each of the n-grams within the word (e.g. for a four letter word there are two trigrams, three bigrams and four monograms) were summed together. Table 1 shows the mean monogram, bigram and trigram summed frequencies for each of the conditions. The table shows that n-gram frequencies in the two orthographically familiar conditions were the same (or even slightly higher in the infrequent condition). For the infrequent words, n-gram frequencies were always significantly lower in the orthographically unfamiliar compared to the orthographically familiar condition. The initial trigram is most important in preprocessing words (Rayner, Well, Pollatsek, & Bertera, 1982). In line with the summed n-gram data, position specific token initial trigram frequencies of the critical words were significantly smaller for the unfamiliar ($M = 166$, $SD = 303$) compared to the familiar ($M = 1307$, $SD = 3087$) infrequent words, $t(39) = 2.28$, $p < .05$, whereas there was no difference for the familiar frequent ($M = 992$, $SD = 903$) and infrequent ($M = 1307$) words ($t < 1$). Type frequencies followed similar patterns to token frequencies for all of the measures described above.

(Insert Table 1 here)

Orthographic familiarity is also associated with differences in the number of orthographic neighbours. The number of orthographic neighours was calculated using the entire English Lexicon Project database (Balota et al., 2002). There was no difference in the number of orthographic neighbours between the frequent ($M = 7.5$, $SD = 4.5$) and infrequent ($M = 8.1$, $SD = 4.8$) orthographically familiar words ($t < 1$) but there were more orthographic neighbours for the orthographically familiar ($M = 8.1$) than the

orthographically unfamiliar ($M = 2.2$, $SD = 2.7$) infrequent words, $t(38) = 7.37$, $p < .001$. The number of higher frequency neighbours for each of the items was also calculated for the frequent and infrequent orthographically familiar words (as noted above, the orthographically unfamiliar infrequent words had very few neighbours). The infrequent orthographically familiar words had significantly more higher frequency neighbours ($M = 4.8$, $SD = 3.8$) than the frequent words ($M = 0.7$, $SD = 0.8$), $t(38) = 6.66$, $p < .001$. These differences reflect the fact that the frequent words are so highly frequent that there are fewer words that are more frequent than them, compared to the infrequent words.[2]

In order to ensure that the critical words were not predictable within the context of the sentence, sentence completion norms were obtained. Twelve participants were given the beginning portions of the sentence up to the critical word, and were asked to provide a single word that they felt could fit as the next word in the sentence. Of all of the completions, only two (0.4%) were correct. Therefore none of the critical words were predictable from the sentence context.

There were 39 critical words in each condition, all were either four or five letters long and they were matched for length across the three conditions with a mean word length of 4.5 characters ($SD = 0.5$). Each set of critical words was embedded in the same neutral sentential frame up to and including the word after the critical word. Each of the sentences was no longer than one line of text (80 characters). The word preceding the critical word (word n-1) was either five or six letters long with a mean word frequency of 147 counts per million ($SD = 170$). A full list of materials is provided in the Appendix.

Three lists of 123 sentences were constructed and ten participants were randomly allocated to each list. Each list included all 117 experimental sentences with an additional

six practice sentences at the beginning. For each third of the lists, the conditions were rotated across the three lists[3]. The order of the items within each third of the list was different, but overall the sentences were presented in a fixed pseudo-random order, ensuring that repeated items were widely distributed throughout the experiment. Thirty-eight of the sentences were followed by a comprehension question.

*Procedure.* Eye movements were monitored using a Dual Purkinje Image eye tracker. Viewing was binocular but only the movements of one eye were monitored (the right eye was monitored for 18 participants and the left eye was monitored for 12 participants[4]). The sentences were presented in light cyan on a black background with characters presented in courier font. The viewing distance was 80cm and 3.7 characters subtended one degree of visual angle. The resolution of the eye tracker is 10 min of arc and the sampling rate was every millisecond.

Participants were instructed to understand the sentences to the best of their ability. A bite bar and head restraint were used to minimize head movements. The accuracy of the eye-tracker was checked (and re-calibrated when necessary) before each trial. After each sentence, participants pressed a button box to continue and to respond yes/no to comprehension questions. The entire experiment lasted approximately 30 minutes.

*Analyses.* Fixations shorter than 80ms that were within one character of the next or previous fixation were incorporated into that fixation. Any remaining fixations shorter than 80ms and longer than 1200ms were discarded. 3.7 percent of trials were excluded due to either no first pass fixations on the sentence prior to word n-1, or tracker loss or blinks on first pass reading of word n-1 or the critical word.

Results

The influence of orthographic familiarity and word frequency on saccade programming was assessed by fixation durations on, and the probability of skipping, the critical word. The duration of the first fixation, single fixation durations, gaze duration (the sum of fixations on a word before leaving it), and total time (the sum of all fixations within a word) were calculated. Parafoveal-on-foveal effects were assessed by examining the duration of the fixation directly before fixating the critical word. In addition, analyses of saccade targeting to the critical word and the effect of the critical word on interword regressions are reported.

Repeated measures Analyses of Variance (ANOVAs) were undertaken across the three conditions for both participants' ($F_1$) and items' ($F_2$) means. Note that the items are the sentences, such that the sentence beginnings were identical for each item, and the three word conditions were manipulated using a repeated measures design. Such a design not only provides control over sentence context across the conditions, but it also enables the items statistical analyses to be undertaken with the more powerful repeated measures procedures[5]. For cases in which the ANOVAs showed a main effect across the three conditions, paired samples t-tests were undertaken. Comparisons between the frequent and infrequent orthographically familiar conditions were used to test for an effect of word frequency. Comparisons between the orthographically familiar and unfamiliar infrequent conditions were used to test for an effect of orthographic familiarity. The mean error rate on the comprehension questions was four percent, indicating that participants properly read and understood the sentences; trials were included regardless of question responses.

*Parafoveal-on-foveal effects.* Table 2 shows mean reading times on word n-1 and fixation durations prior to fixating the critical word. There were no effects of the condition of the critical word on first fixation durations, single fixation durations or gaze durations on word n-1 ($Fs < 1$). Due to acuity limitations, the characteristics of a parafoveal word are only likely to influence processing on previous fixations for saccades launched from near launch sites (Kennison & Clifton, 1995; Lavigne, Vitu, & d'Ydewalle, 2000; Rayner, 1975; Rayner, Binder, Ashby, & Pollatsek, 2001; White & Liversedge, 2006b). Therefore influences of the critical word on the fixation duration prior to fixating it were calculated for all of the data and only for the 88% of cases in which saccades were launched from word n-1.

(Insert Table 2 here)

There was a main effect of condition on the duration of the fixation prior to fixating the critical word for all data, $F_1 (2,58) = 3.24$, $MSe = 174$, $p < 0.05$, partial $\eta^2 = 0.101$, $F_2 (2,76) = 3.19$, $MSe = 148$, $p < 0.05$, partial $\eta^2 = 0.077$, and for saccades launched from word n-1, $F_1 (2,58) = 4.55$, $MSe = 162$, $p < 0.05$, partial $\eta^2 = 0.136$, $F_2 (2,76) = 3.13$, $MSe = 174$, $p < 0.05$, partial $\eta^2 = 0.076$. For infrequent critical words, there was a small (6ms) parafoveal-on-foveal effect such that prior fixation durations were longer when the critical word was orthographically unfamiliar compared to when it was orthographically familiar and these effects were significant for saccades launched from word n-1, $t_1 (29) = 2.37$, $p < 0.05$, $t_2 (38) = 2.49$, $p < 0.05$, but significant only across items, $t_2 (38) = 2.33$, $p < 0.05$, and not participants, $t_1 (29) = 1.85$, $p = 0.075$, for all of the data. In contrast, for both

sets of analyses for orthographically familiar words there was no difference in prior

fixation duration between the frequent and infrequent conditions ($ts < 1$).[6]

It has been suggested that parafoveal-on-foveal effects can sometimes occur as a

result of inaccurate saccade targeting (Drieghe et al., 2007; Rayner, Warren et al., 2004;

Rayner, White et al., 2003). However, if the parafoveal-on-foveal effects reported above

had arisen due to mistargeting of saccades, then there perhaps also should have been a

parafoveal-on-foveal effect of word frequency. The following analysis aims to identify if

parafoveal-on-foveal effects occurred when cases in which there were most likely to have

been mistargeted saccades were eliminated. Saccades that were intended to land on the

critical word but that undershot were most likely to have landed on the three final

characters of word n-1. For fixations launched from word n-1, but not from the three final

characters of word n-1, the main effect of condition was marginal, $F_1 (2,58) = 2.84$, $MSe = $

223, $p = 0.067$, partial $\eta^2 = 0.089$, $F_2 (2,76) = 2.7$, $MSe = 266$, $p = 0.074$, partial $\eta^2 = $

0.066. Although there was no effect of parafoveal word frequency on these fixation

durations ($ts < 1$), results indicate that prior fixation durations in this region were longer

prior to orthographically unfamiliar compared to orthographically familiar infrequent

words, $t_1 (29) = 2.22$, $p < 0.05$, $t_2 (38) = 2.74$, $p < 0.01$.

Together, the absence of a parafoveal-on-foveal effect of word frequency, and an

indication of a parafoveal-on-foveal effect of orthography more than three characters from

the critical words, suggests that the orthographic parafoveal-on-foveal effects shown here

may not simply be explained by inaccurate saccade targeting (see Starr & Inhoff, 2004 for

a similar finding). Importantly, the present study may provide very favourable conditions

for parafoveal-on-foveal effects to arise. That is, both the foveal (word n-1) and parafoveal

(critical) words were short and the orthographic familiarity of the entire parafoveal word was strongly manipulated. (Though note that it could be that only the familiarity of some of the letters, such as those at the word beginning, could have produced the effect). Consequently, the parafoveal-on-foveal effects shown here may not hold under less favourable conditions, for example, when there are long parafoveal words or only the orthographic familiarity of the word beginning is manipulated (White & Liversedge, 2004, 2006b). Critically, despite the favourable conditions, and consistent with prior research (Rayner et al., 1998), there was no effect of word frequency on prior fixation durations for orthographically familiar words. [7]

*Reading measures for the critical word.* The mean reading times, refixation probabilities and spillover fixation durations for the critical word are shown in Table 3. There were significant main effects of condition for all of the reading measures on the critical word ($Fs > 15$, $ps < .001$). For orthographically familiar words, reading times for the critical word were shorter on frequent than infrequent words for first fixation durations, $t_1 (29) = 8.96$, $p < 0.001$, $t_2 (38) = 5.34$, $p < 0.001$, single fixation durations, $t_1 (29) = 8.01$, $p < 0.001$, $t_2 (38) = 6.01$, $p < 0.001$, gaze durations, $t_1 (29) = 10.18$, $p < 0.001$, $t_2 (38) = 7.25$, $p < 0.001$, and total time, $t_1 (29) = 8.8$, $p < 0.001$, $t_2 (38) = 5.98$, $p < 0.001$. There were also significantly more refixations on the infrequent, compared to frequent, orthographically familiar words, $t_1 (29) = 4.6$, $p < 0.001$, $t_2 (38) = 4.05$, $p < 0.001$. These results clearly demonstrate that even when orthographic familiarity is carefully controlled, the lexical variable of word frequency influences word processing, even for the first fixation on the word.

(Insert Table 3 here)


For infrequent words, reading times for the critical word were consistently numerically longer, and there were numerically more refixations, on the orthographically unfamiliar than the orthographically familiar critical words, though these effects were much smaller than those for word frequency. The effect of orthographic familiarity was significant for single fixation durations, $t_1 (29) = 2.27$, $p < 0.05$, $t_2 (38) = 2.46$, $p < 0.05$; significant across participants, $t_1 (29) = 2.02$, $p = 0.05$, but not items, $t_2 (38) = 1.5$, $p = 0.143$, for first fixation durations; significant across participants, $t_1 (29) = 2.87$, $p < 0.01$, marginal across items, $t_2 (38) = 1.93$, $p = 0.061$, for gaze durations; significant across participants, $t_1 (29) = 2.54$, $p < 0.05$, but not items, $t_2 (38) = 1.7$, $p = 0.098$, for refixation probability, and there was no effect for total time, $t_1 (29) = 1.4$, $p = 0.171$, $t_2 < 1$. The direction of these orthographic familiarity effects is consistent with previous evidence from sentence reading (Lima & Inhoff, 1985), and consistent with some (Biederman, 1966; Seidenberg et al., 1984; Waters & Seidenberg, 1985) but not other (Broadbent & Gregory, 1968; Rice & Robinson, 1975) studies using isolated word processing tasks.

There was also a main effect of condition on the duration of the spillover fixation, after leaving the critical word either to the right or left, $F_1 (2,58) = 3.83$, $MSe = 209$, $p < 0.05$, partial $\eta^2 = 0.117$, $F_2 (2,76) = 3.68$, $MSe = 299$, $p < 0.05$, partial $\eta^2 = 0.088$. For orthographically familiar words, fixations were longer following fixation of the infrequent compared to the frequent words, $t_1 (29) = 2.44$, $p < 0.05$, $t_2 (38) = 2.75$, $p < 0.01$. In contrast, for the infrequent words, there was no difference in the duration of the fixation following the orthographically unfamiliar compared to the familiar words, $t_1 < 1$, $t_2 (38) =$

1.24, $p = 0.223$. The results are consistent with previous studies showing spillover effects of word frequency (Kennison & Clifton, 1995; Kliegl et al., 2006; Rayner & Duffy, 1986), though note that such effects do not always occur (White & Liversedge, 2006a).

The distribution of single fixation durations on the critical word is shown in Figure 1. Note that the longer fixation durations in the infrequent, compared to frequent word conditions is characterised by a rightward shift in the distribution, consistent with previous studies (Rayner, 1995; Rayner, Liversedge et al., 2003). Numerical differences in the distributions between the frequent and infrequent orthographically familiar conditions occur no earlier than differences in the distributions between the orthographically familiar and unfamiliar infrequent word conditions. These numerical patterns indicate that the effect of word frequency during single fixation durations occurs no later than that of orthographic familiarity.

Overall, the results show that word frequency has a robust and long lasting effect on word processing. In contrast, orthographic familiarity has a numerically smaller influence for all of the reading time measures on the critical word. The absence of significant orthographic familiarity effects in later measures such as spillover, indicates that orthographic familiarity exerts a relatively small and short-lived influence on word processing compared to effects of word frequency.


(Insert Figure 1 here)


*Critical word skipping probabilities*. Table 4 shows the probability of skipping the critical word in each of the conditions. Due to acuity limitations, effects of word skipping

are calculated for all of the data and only for the 88% of cases in which saccades were launched from word n-1. There was a main effect of condition on the probability of skipping the critical word both for all of the data, $F_1 (2,58) = 3.2$, $MSe = 0.014$, $p < 0.05$, partial $\eta^2 = 0.099$, $F_2 (2,76) = 3.8$, $MSe = 0.019$, $p < 0.05$, partial $\eta^2 = 0.091$, and for saccades launched from word n-1, $F_1 (2,58) = 5.07$, $MSe = 0.006$, $p < 0.01$, partial $\eta^2 = 0.149$, $F_2 (2,76) = 5.55$, $MSe = 0.006$, $p < 0.01$, partial $\eta^2 = 0.127$. For all of the data there were no significant effects of either word frequency, $t_1 (29) = 1.79$, $p = 0.084$, $t_2 (38) = 1.83$, $p = 0.075$, or orthographic familiarity ($t$s < 1) on word skipping. For saccades launched from word n-1 there was no effect of orthographic familiarity on word skipping for the infrequent words ($t$s < 1). However for saccades launched from word n-1, for orthographically familiar critical words, frequent words were more likely to be skipped than infrequent words, $t_1 (29) = 2.19$, $p < 0.05$, $t_2 (38) = 2.17$, $p < 0.05$.

(Insert Table 4 here)

Compared to the results for saccades launched from word n-1, a similar pattern of effects holds for saccades launched from three or less characters from the critical word. There was a main effect of condition, $F_1 (2,58) = 7.01$, $MSe = 0.019$, $p < 0.01$, partial $\eta^2 = 0.195$, $F_2 (2,76) = 6.74$, $MSe = 0.017$, $p < 0.01$, partial $\eta^2 = 0.151$, which was characterised by a greater probability of skipping frequent compared to infrequent orthographically familiar words, $t_1 (29) = 2.22$, $p < 0.05$, $t_2 (38) = 2.08$, $p < 0.05$. However there was no significant difference in the probability of skipping orthographically familiar compared to orthographically unfamiliar infrequent words, $t_1 (29) = 1.78$, $p = 0.086$, $t_2 (38) = 1.31$, $p =$

0.2. These results show that word frequency, but not orthographic familiarity, significantly influence saccade programming such that this determines whether words are skipped[8]. Critically, the effect of word frequency on word skipping indicates that the lexical characteristics of words are processed in the parafovea and that they can have an early influence on word processing and eye movement behaviour. Furthermore, the fact that orthographic familiarity produces a parafoveal-on-foveal effect, but does not influence word skipping, indicates that although these two measures both reflect early parafoveal processing of words, they may be determined by qualitatively different aspects of eye movement control.

*Saccade metrics for the initial first pass fixation on the critical word.* Mean landing positions, launch positions and saccade lengths for first pass saccades launched from word n-1 to the critical word are shown in Table 4. There were no effects of condition on landing positions ($F$s < 1), launch positions, $F_1$ (2,58) = 2.25, $MSe$ = 0.174, $p$ = 0.114, partial $\eta^2$ = 0.072, $F_2$ (2,76) = 1.87, $MSe$ = 0.095, $p$ = 0.162, partial $\eta^2$ = 0.047, or saccade lengths, $F_1$ (2,58) = 1.88, $MSe$ = 0.16, $p$ = 0.162, partial $\eta^2$ = 0.061, $F_2$ (2,76) = 2.25, $MSe$ = 0.09, $p$ = 0.113, partial $\eta^2$ = 0.056.

These findings suggest that neither the orthographic familiarity nor word frequency of four and five letter words influences saccade targeting to words. These results contrast with evidence showing that orthographic familiarity does influence where longer words are first fixated (Hyönä, 1995; Radach et al., 2004; Vonk, et al. 2000; White & Liversedge, 2004, 2006a, 2006b). However, note that initial trigrams used in the experiments with longer words were often much more infrequent than those used here. Therefore it is possible that the orthographic familiarity of shorter words may influence

saccade targeting with much stronger manipulations of orthographic familiarity. It is also

possible that orthographic influences on saccade targeting, have no, or much smaller,

effects for shorter words.

*Regressions.* Regression probabilities out of and in to the critical word are

presented in Table 4. There was no effect of condition on the probability of making a first

pass regression out of the critical word ($F$s < 1). The probability of making a regression in

to the critical word was investigated as a function of whether the critical word was fixated

on first pass[9]. A 2 (skip vs fixate) X 3 (condition) ANOVA showed a significant main

effect of skipping, $F_1(1,27) = 70.36$, $MSe = 0.054$, $p < 0.001$, partial $\eta^2 = 0.723$, $F_2(1,37)$

$= 194.14$, $MSe = 0.031$, $p < 0.001$, partial $\eta^2 = 0.84$, such that there were more regressions

into the critical word when it was skipped (0.41) compared to when it was fixated (0.11)

on first pass. There was also a main effect of condition, $F_1(2,54) = 9.92$, $MSe = 0.032$, $p <$

0.001, partial $\eta^2 = 0.269$, $F_2(2,74) = 10.85$, $MSe = 0.049$, $p < 0.001$, partial $\eta^2 = 0.227$, and

an interaction, $F_1(2,54) = 7.07$, $MSe = 0.029$, $p < 0.01$, partial $\eta^2 = 0.208$, $F_2(2,74) = 7.86$,

$MSe = 0.042$, $p < 0.01$, partial $\eta^2 = 0.175$.

For cases in which the critical word was fixated on first pass there was no effect of

condition on the probability of making a regression into the critical word, $F_1(2,58) = 2.16$,

$MSe = 0.004$, $p = 0.124$, partial $\eta^2 = 0.069$, $F_2(2,76) = 1.65$, $MSe = 0.006$, $p = 0.199$,

partial $\eta^2 = 0.042$. In contrast, for cases in which the critical word was skipped on first

pass there was a main effect of condition on the proportion of regressions made into the

critical word, $F_1(2,54) = 9.09$, $MSe = 0.056$, $p < 0.001$, partial $\eta^2 = 0.252$, $F_2(2,74) =$

10.01, $MSe = 0.085$, $p < 0.001$, partial $\eta^2 = 0.213$. For cases in which the critical word was

skipped on first pass, for orthographically familiar words there were more regressions into

infrequent than frequent critical words, $t_1$ (27) = 2.92, $p < 0.01$, $t_2$ (37) = 3.5, $p < 0.01$.

However there was no difference in the proportion of regressions made into the critical

word for infrequent orthographically unfamiliar compared to familiar words ($t$s < 1.1).

The finding that word frequency influences the probability of making a regression back to

a skipped word supports results reported by Vitu and McConkie (2000). It is particularly

interesting that the present findings show that word frequency, but not orthographic

familiarity, significantly influenced regressions. The results suggest that the linguistic

influence on regressions is at a lexical, rather than sub-lexical, level. The effect of lexical

processing difficulty on regressions supports the suggestion made by Engbert et al. (2005)

that inter-word regressions can be caused by incomplete lexical access. However note that

processing difficulty related to post-lexical semantic integration may also be related to

such an effect of word frequency. The findings are also consistent with previous research

suggesting that regressive saccades can target areas of processing difficulty quite

accurately (Frazier & Rayner, 1982; Kennedy, Brooks, Flynn, & Prophet, 2003;

Meseguer, Carreiras, & Clifton, 2002).

## Discussion

The results clearly show that word frequency, independent of orthographic

familiarity, influences the probability of skipping words and fixation durations on words.

However there were no effects of word frequency on prior fixation durations. In contrast,

orthographic familiarity had a small effect on prior fixation durations and fixation

durations on words, but no influence on the probability of word skipping. The key

implications are therefore that lexical processing of fixated words can influence saccade

programming as shown by fixation durations, and that lexical processing of parafoveal words can influence saccade programming as shown by word skipping. The results have important implications for models of eye movement control in reading and these are discussed below. In addition, there was a small orthographic parafoveal-on-foveal effect but no evidence of lexical parafoveal-on-foveal effects. The implications of these findings are considered at the end of the Discussion.

*Saccade programming: Fixation durations and word targeting*

The findings reported here demonstrate that models of eye movement control in reading should include a saccade programming mechanism which is sensitive to lexical (or even post-lexical) processing of word frequency. Therefore, there must be sufficient time during a fixation for both lexical processing and saccade programming to occur, before the metrics and timing of the subsequent saccade are finalised. The present findings are particularly crucial for models in which linguistic factors modulate when saccade programming is initiated because, due to the time required for saccade programming, such linguistic influences may have to occur early during fixations (Engbert et al., 2002, 2005; Kliegl & Engbert, 2003; Pollatsek et al., 2006; Reichle et al., 1998, 1999, 2003, 2006; Richter et al., 2006).

In the E-Z reader model (Pollatsek et al., 2006; Reichle et al., 1998, 1999, 2003, 2006) completion of an initial stage of word processing ($L_1$) triggers the initiation of a saccade programme, which subsequently takes approximately 125ms to complete. Therefore the time at which this first stage of word processing is completed effectively produces the linguistic influence on the duration of fixations. In order for a word to be skipped, both the fixated word and the parafoveal word would have to have reached the $L_1$

stage before the skipping programme could be finalised. In the SWIFT model, a random stochastic process influences when saccade programming begins, and this is delayed or cancelled by linguistic processing using a mechanism referred to as foveal inhibition (Engbert et al., 2002, 2005; Kliegl & Engbert, 2003; Richter et al., 2006). Therefore similar to the E-Z reader model, in the SWIFT model linguistic influences on fixation durations must also occur prior to saccade programming. Crucially, in order for such accounts to explain the results reported here, the lexical variable of word frequency must have an influence early during a fixation before saccade programming commences. For example, for a fixation lasting 250ms with 125ms of saccade programming, lexical influences on saccade programming would have to occur in the initial 125ms of the fixation. Given such an account, either lexical processing must be very fast, or a significant amount of this processing may be undertaken on previous fixations (Findlay & White, 2003).

An alternative possibility is that linguistic factors may influence the saccade programme later during a fixation, but before the non-labile stage of saccade programming. For example in the SWIFT model, determination of the saccade target (which word is fixated) occurs at the end of the labile stage of saccade programming. If linguistic processing could influence saccade programming during the labile stage then there would be more time during the fixation for lexical processing to occur such that it could then influence when or where the eyes moved. For example, for a fixation lasting 250ms with 50ms for the non-labile stage of saccade programming, 200ms of linguistic processing time may be available before the linguistic influence on the saccade programme is finalised at the end of the labile stage.

In the E-Z reader model, once the first stage of word processing is complete ($L_1$) and saccade programming has been initiated, a second stage of word processing ($L_2$) is achieved before attention shifts to the following word (Pollatsek et al., 2006; Reichle et al., 1998, 1999, 2003, 2006). As saccade programming time is relatively constant, the duration of $L_2$ determines the time between the attention shift and when the eyes move, during which the following word can be preprocessed. For the E-Z reader model, the influence of $L_2$ on the degree of parafoveal processing is crucial to explaining both spillover effects (e.g. Rayner & Duffy, 1986) and modulation of parafoveal processing by foveal load (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; White, Rayner, & Liversedge, 2005). If $L_1$ is of a lexical nature, this raises the question of exactly how $L_2$ is different to $L_1$. Future models based on multiple word processing stages may need to specify the types of processing entailed in each stage more precisely. For a more detailed discussion of how even post-lexical factors may possibly influence the first stage of word processing, see Reichle, Pollatsek, and Rayner (2007).

The present study also raises the issue of precisely how the orthographic familiarity of words influences fixation durations on those words. In the E-Z reader model (Pollatsek et al., 2006; Reichle et al., 1998, 1999, 2003, 2006) it could be argued that orthographic familiarity may influence the initial stage of word processing ($L_1$) which influences the current fixation duration, but not the second stage of word processing ($L_2$) which can influence subsequent fixation durations. Such a suggestion would be consistent with the finding that orthographic familiarity had a short-lived effect on reading times on the critical word in the present study (for similar reasoning see Reingold, 2003; Reingold & Rayner, 2006).

The results also showed that the lexical characteristics of words influenced the probability of making a regression back to words if they were skipped on first pass. In line with these findings, the SWIFT model of eye movement control is unique in that it predicts that regressions are directed to words that are not completely processed on first pass (Engbert et al. 2002, 2005; Kliegl & Engbert, 2003; Richter et al., 2006) (see also: Reilly & Radach, 2003, 2006). It seems especially noteworthy that orthographic familiarity did not significantly influence word skipping or regression probabilities, whilst word frequency did influence these measures. These results suggest that lexical factors have a more critical role than sub-lexical factors in determining which words are fixated and refixated. However, note that although lexical factors have been shown here to influence both when and where the eyes move, research suggests that these two processes may in fact be controlled in qualitatively different ways. For example, whilst fixation durations are influenced by both foveal and parafoveal linguistic information, word skipping may be influenced by linguistic information only in the parafovea (White, 2007, see also Drieghe, Rayner, & Pollatsek, 2005).

*Parafoveal-on-foveal effects*

The present study also has important implications for the contentious issue of parafoveal-on-foveal effects. The results showed that, at least when both foveal and parafoveal words are short and the orthographic familiarity of the entire parafoveal word is manipulated, the orthographic characteristics of the critical word had a small (6 ms) influence on prior fixation durations. These results are in line with previous sentence reading studies showing parafoveal-on-foveal effects at least at the level of orthographic familiarity (Inhoff et al., 2000; Pynte et al., 2004; Rayner, 1975; Starr & Inhoff, 2004;

Underwood et al., 2000). It could be that only the orthographic familiarity of the initial letters of the word is critical. Furthermore, other characteristics associated with orthographic familiarity, such as the informativeness of the component letter sequences, may have driven the effects. Fixations are often mislocated (McConkie et al., 1988; Nuthmann et al., 2005) and some parafoveal-on-foveal effects may be explained by such mislocated fixations (Dreighe et al., 2005, 2007; Rayner, Warren et al., 2004; Rayner, White et al., 2003). However the present results indicate that the orthographic parafoveal-on-foveal effects shown here are not simply isolated to fixations three or less characters from the critical word. Therefore, either a substantial proportion of mislocated fixations must land more than three characters from the critical word or, as concluded by Starr and Inhoff (2004), there are parafoveal-on-foveal effects of orthography that are not caused by mislocated fixations.

Given that the explanation of mislocated fixations does not seem feasible in this case, there are two alternative accounts for the small orthographic parafoveal-on-foveal effects shown here. First, attention may be allocated to multiple words in parallel such that processing of the foveal word may occur simultaneously with processing of the orthographic characteristics of the parafoveal word (Engbert et al., 2002, 2005; Kennedy, 2000; Kliegl & Engbert, 2003; McDonald et al., 2005; Reilly & Radach, 2003, 2006; Richter et al., 2006). Second, attention may be allocated to words one at a time (as in serial models) but the orthographic characteristics of the parafoveal words may be concurrently processed in a manner that does not require attention. Such early pre-attentive visual processing of the parafoveal word may impact on processing of the foveal word in order to generate orthographic parafoveal-on-foveal effects (Pollatsek et al., 2006;

Reichle et al., 2003, 2006). Note that in the E-Z reader model, pre-attentive processes that may induce parafoveal-on-foveal effects are quite separate to the linguistic processes associated with the $L_1$ and $L_2$ stages of word recognition, that have a more primary role in influencing saccade programming during reading.

To be clear, the parafoveal-on-foveal effects shown here were not necessarily mediated by the same attention based processes that are generally allocated for reading of words. That is, parafoveal-on-foveal effects may not necessarily impact on the reading process (see Drieghe et al., 2007). Therefore although there may be temporal overlap between processing of the foveal word and processing related to the orthographic characteristics of the parafoveal word, these separate processes may be independent and perhaps even qualitatively different. For example, the orthographic parafoveal-on-foveal effects may be mediated by visual, not linguistic, processes. Overall, the orthographic parafoveal-on-foveal effects reported here can not distinguish between the parallel attention (Engbert et al., 2002, 2005; Kennedy, 2000; Kliegl & Engbert, 2003; McDonald et al., 2005; Reilly & Radach, 2003, 2006; Richter et al., 2006) and pre-attentive (Pollatsek et al., 2006; Reichle et al., 2003, 2006) explanations provided by the models.

The present study showed no effect of word frequency on prior fixation durations. This is despite the study providing favourable conditions for lexical parafoveal-on-foveal effects to occur by using short foveal and parafoveal words. Note that the foveal word n-1 were frequent words (see Method) so there should have been sufficient time or processing resources to process the parafoveal word (Henderson & Ferreira, 1990; Kennison & Clifton, 1995; White et al., 2005). The results are inline with previous experimental studies showing no evidence of lexical parafoveal-on-foveal effects (Altarriba et al., 2001;

Calvo & Meseguer, 2002; Henderson & Ferreira, 1993; Inhoff et al., 2000; Rayner, 1975; Rayner et al., 1998; Rayner, Juhasz et al., 2007; Schroyens et al., 1999), but they contrast with studies that have shown lexical parafoveal-on-foveal effects using very large data sets across a corpus of data or a combination of different studies (Kennedy & Pynte, 2005; Kliegl et al., 2006; Pynte & Kennedy, 2006). It is not clear why standard experimental studies consistently show such a different pattern of results to studies based on larger datasets. At the very least, the data suggest that if lexical parafoveal-on-foveal effects do occur then these must be of a small or specific nature, such that they are so elusive in experimental studies. Given that experimental studies provide much greater control over other variables than corpus based studies, it is important that findings that are shown in corpus studies, but not in sentence based experimental studies, be treated with caution (Rayner, Pollatsek, Drieghe, Slattery, & Reichle, 2007). Overall, perhaps in the vast majority of cases the lexical characteristics of words are processed serially (Morrison, 1984; Pollatsek et al., 2006; Reichle et al., 1998, 1999, 2003, 2006) but with very large samples of data it can be shown that lexical parafoveal-on-foveal effects do occur, though it is still not entirely clear whether such effects are due to parallel processing of words, mislocated fixations or other variables.

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.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database*. [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Balota, D.A., Cortese, M.J., Hutchison, K.A., Neely, J.H., Nelson, D., Simpson, G.B., & Treiman, R. (2002). *The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and nonwords.* http://elexicon.wustl.edu/, Washington University.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19,* 967-983.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory & Language,* 48, 615-634.

Biederman, G.B. (1966). The recognition of tachistoscopically presented five-letter words as a function of digram frequency. *Journal of Verbal Learning and Verbal Behaviour, 5,* 208-209.

Broadbent, D.E., & Gregory, M. (1968). Visual perception of words differing in letter digram frequency. *Journal of Verbal Learning & Verbal Behavior,* 7, 569-571.

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

Brysbaert, M., & Vitu, F. (1998). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception* (pp. 125-148). Oxford, UK: Elsevier.

Calvo, M.G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical, and contextual factors. *The Spanish Journal of Psychology, 5,* 66-77.

Deubel, H., O'Regan, J.K., & Radach, R. (2000). Attention, information processing and eye movement control. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 355-374). Oxford, UK: Elsevier.

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.

Drieghe, D., Rayner, K., & Pollatsek, A. (2007). Mislocated fixations can account for parafoveal-on-foveal effects in eye movements during reading. *Quarterly Journal of Experimental Psychology, in press.*

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.

Engbert, R., Nuthmann, A., Richter, E.M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112,* 777-813.

Feng, G. (2006). Eye movements as time-series random variables: A stochastic model of eye movement control in reading. *Cognitive Systems Research, 7,* 70-95.

Findlay, J.M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioural and Brain Sciences, 22*, 661-674.

Findlay, J.M., & White, S.J. (2003). Serial programming for saccades: Does it all add up? *Behavioral & Brain Sciences,* 26, 483-484.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the Analysis of Structurally Ambiguous Sentences. *Cognitive Psychology, 14,* 178-210.

Gernsbacher, M.A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General, 113,* 256-281.

Grainger, J., & Dijkstra, T. (1996). Visual word recognition: Models and experiments. In: T. Dijkstra, & K. de Smedt (Eds.) *Computational psycholinguistics: AI and connectionist models of human language processing.* London: Taylor & Francis.

Henderson, L. (1982). *Orthography and word recognition in reading.* London: Academic Press.

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.

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.

Holmes, V.M., & O'Regan, J.K. (1987). Decomposing French words. In J.K. O'Regan & A. Lévy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 459-466). Amsterdam: Elsevier.

Humphreys, G.W., Evett, L.J., & Quinlan, P.T. (1990). Orthographic processing in visual word recognition. *Cognitive Psychology, 22,* 517-560.

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

Hyönä, J., & Bertram, R. (2004). Do frequency characteristics of non-fixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology, 16,* 104-127.

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.

Inhoff, A.W., & Eiter, B.M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 979-995.

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). Oxford, UK: Elsevier.

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

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

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

Just, M.A., & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329-354.

Kennedy, A. (1998). The influence of parafoveal words on foveal inspection time: Evidence for a processing trade off. In G. Underwood (Ed.), *Eye Guidance in reading and scene perception* (pp. 149-179). Oxford, UK: Elsevier.

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology, 53A,* 429-455.

Kennedy, A., Brooks, R., Flynn, L.-A., & Prophet, C. (2003). The reader's spatial code. In J. Hyönä, R. Radach, & H. Deubel (Eds). *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 193-212). Amsterdam: Elsevier.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research,* 45, 153-168.

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology, 55A,* 1307-1337.

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

Kliegl, R., & Engbert, R. (2003). SWIFT explorations. In J. Hyönä, R. Radach, & H. Deubel (Eds). *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 391-411). Amsterdam: Elsevier.

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review, 12,* 132-138.

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.

Lavigne, F., Vitu, F., & d'Ydewalle, G. (2000). The influence of semantic context on initial eye landing sites in words. *Acta Psychologica, 104,* 191-214.

Legge, G.E., Klitz, T.S., & Tjan, B.S. (1997). Mr. Chips: An ideal-observer model of reading. *Psychological Review, 104,* 524-553.

Legge, G.E., Hooven, T.A., Klitz, T.S., Mansfield, J.S., & Tjan, B.S. (2002). Mr. Chips 2002: new insights from an ideal-observer model of reading. *Vision Research, 42,* 2219-2234.

Lima, S.D., & Inhoff, A.W. (1985). Lexical access during eye fixations in reading: Effects of word initial letter sequence. *Journal of Experimental Psychology: Human Perception & Performance,* 11, 272-285.

Liversedge, S.P., Rayner, K., White, S.J., Vergilino-Perez, D., Findlay, J.M., & Kentridge, R.W. (2004). Eye movements when reading disappearing text: is there a gap effect in reading? *Vision Research, 44,* 1013, 1024.

Liversedge, S.P., & White, S.J. (2003). Psycholinguistic processes affect fixation durations and orthographic information affects fixation locations: Can E-Z Reader cope? *Behavioral and Brain Sciences, 26,* 492-493.

Martin, G.L. (2004). Encoder: A connectionist model of how learning to visually encode fixated text images improves reading fluency. *Psychological Review, 111,* 617-639.

Mayzner, M.S., & Tresselt, M.E. (1965). Tables of single-letter and digram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements, 1,* 13-32.

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.

McConkie, G.W., & Yang, S-N., (2003). How cognition affects eye movements during reading. In J. Hyönä, R. Radach, & H. Deubel (Eds). *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 413-427). Amsterdam: Elsevier.

McDonald, S.A., Carpenter, R.H.S., & Shillcock, R.C., (2005). An anatomically constrained, stochastic model of eye movement control in reading. *Psychological Review, 112,* 814-840.

Meseguer, E., Carreiras, M., & Clifton, C. (2002). Overt reanalysis strategies and eye movements during the reading of mild garden path sentences. *Memory & Cognition, 30,* 551-561.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner, & G.W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp.148-197). Hillsdale, NJ: Lawrence Erlbaum.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception & Performance,* 10, 667-682.

Murray, W.S. (1998). Parafoveal pragmatics. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 181-200). Oxford, England: Elsevier.

Nazir, T.A. (2000). Traces of print along the visual pathway. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 3-22). Oxford, UK: Elsevier.

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

O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye movments and their role in visual and cognitive processes* (pp. 395-453). Amsterdam: Elsevier.

O'Regan, J.K. (1992). Optimal viewing position in words and the strategy-tactics theory of eye movements in reading. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 333-354). New York: Springer-Valag.

Perea, M., & Pollatsek, A. (1998). The effects of neighbourhood frequency in reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 767-779.

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.

Pollatsek, A., Reichle, E., & Rayner, K. (2006). Tests of the E-Z reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52,* 1-56.

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

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typographical errors in eye movements in reading. *European Journal of Cognitive Psychology, 16,* 178-202.

Pynte, J., Kennedy, A., & Murray, W.S. (1991). Within-word inspection strategies in continuous reading: Time course of perceptual, lexical and contextual processes. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 458-470.

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

Raney, G.E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology, 49,* 151-173.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65-81.

Rayner, K. (1995). Eye movements and cognitive processes in reading, visual search, and scene perception. In J.M. Findlay, R. Walker, & R.W. Kentridge (Eds.), *Eye movement research: Mechanisms, processes and applications* (pp. 3–12). Amsterdam: North Holland.

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

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E.D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 720-732.

Rayner, K., Binder, K.S., Ashby, J., & Pollatsek, A. (2001). Eye movement control in reading: Word predictability has little influence on initial landing positions in words. *Vision Research, 41,* 943-954.

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

Rayner, K., Fischer, M.H., & Pollatsek, A. (1998). Unspaced text interferes with both word identification and eye movement control. *Vision Research, 38,* 1129-1144.

Rayner, K., & Juhasz, B.J. (2004). Eye Movements in Reading: Old Questions and New Directions. *European Journal of Cognitive Psychology, 16,* 340-352.

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.

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.

Rayner, K., & Morrison, R.M. (1981). Eye movements and identifying words in parafoveal vision. *Bulletin of the Psychonomic Society, 17,* 135-138.

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T.J., & Reichle, E.D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology: General, in press*

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 & Performance,* 22, 1188-1200.

Rayner, K., Slowiaczek, M.L., Clifton, C.Jr., & Bertera, J.H. (1983). Latency of sequential eye movements: implications for reading. *Journal of Experimental Psychology: Human Perception & Performance,* 9, 912-922.

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

Rayner, K., Well, A.D., Pollatsek, A., & Bertera, J.H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics, 31,* 537-550.

Rayner, K., White, S.J., Kambe, G., Miller, B., & Liversedge, S.P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds). *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 213-234). Amsterdam: Elsevier.

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.

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.

Reichle, E. D., Pollatsek, A., & Rayner, K. (2007). Modeling the effects of lexical ambiguity on eye movements during reading. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds). *Eye movements: A window on mind and brain*. Elsevier.

Reichle, E.D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: accounting for initial fixation locations and refixations within the E-Z reader model. *Vision Research,* 39, 4403-4411.

Reichle, E.D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral & Brain Sciences,* 26, 445-526.

Reilly, R.G., & O'Regan, J.K. (1998). Eye movement control during reading: A simulation of some word-targetting strategies. *Vision Research, 38,* 303-317.

Reilly, R.G., & Radach, R. (2003). Foundations of an interactive activation model of eye movement control in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp.429-455). Amsterdam: Elsevier.

Reilly, R.G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7,* 34-55.

Reingold, E.M. (2003). Eye-movement control in reading: Model and predictions. *Behavioral and Brain Sciences, 26,* 500-501.

Reingold, E.M., & Rayner, K. (2006). Examining the word identification stages hypothesised by the E-Z Reader model. *Psychological Science, 17,* 742-746.

Rice, G.A., & Robinson, D.O. (1975). The role of bigram frequency in the perception of words and nonwords. *Memory & Cognition,* 3, 513-518.

Richter, E.M., Engbert, R., & Kliegl, R. (2006). Current advances in SWIFT. *Cognitive Systems Research, 7,* 23-33.

Salvucci, D.D. (2001). An integrated model of eye movements and visual encoding. *Journal of Cognitive Systems Research, 1,* 201-220.

Schiepers, C. (1980). Response latency and accuracy in visual word recognition. *Perception & Psychophysics, 27,* 71-91.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology, 52A,* 1021-1046.

Sears, C.R., Campbell, CR., & Lupker, S.J. (2006). Is there a neighborhood frequency effect in English? Evidence from reading and lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1040-1062.

Seidenberg, M.S., Waters, G.S., Barnes, M.A., & Tanenhaus, M.K. (1984). When does irregular spelling or pronunciation influence word recognition. *Journal of Verbal Learning & Verbal Behaviour, 23,* 383-404.

Starr, M.S., & Inhoff, A.W. (2004). Attention Allocation to the Right and Left of a Fixated Word: Use of Orthographic Information from Multiple Words during Reading. *European Journal of Cognitive Psychology, 16,* 203-225.

Suppes, P. (1990). Eye-movement models for arithmetic and reading performance. In: E. Kowler (Ed.), *Eye Movements and Their Role in Visual and Cognitive Processes* (pp. 455-477). Amsterdam: Elsevier.

Thibadeau, R., Just, M.A., & Carpenter, P.A. (1982). A model of the time course and content of reading. *Cognitive Science, 6,* 157-203.

Underwood, G., Binns, A., & Walker, S. (2000). Attentional demands on the processing of neighbouring words. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 247-268). Oxford, UK: Elsevier.

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

Vonk, W., Radach, R., & van Rijn, H. (2000). Eye guidance and the saliency of word beginnings in reading text. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 269-299). Oxford, UK: Elsevier.

Waters, G.S., & Seidenberg, M.S. (1985). Spelling-sound effects in reading: Time-course and decision criteria. *Memory & Cognition, 13*, 557-572.

White, S.J. (2007). Foveal load and parafoveal processing: The case of word skipping. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds). *Eye movements: A window on mind and brain*. Elsevier.

White, S.J., & Liversedge, S.P. (2004). Orthographic Familiarity Influences Initial Eye Fixation Positions in Reading. *European Journal of Cognitive Psychology, 16*, 52-78.

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.

White, S.J., & Liversedge, S.P. (2006b). Linguistic and non-linguistic influences on the eyes' landing positions during reading. *Quarterly Journal of Experimental Psychology, 59,* 760–782.

White, S.J., Rayner, K., & Liversedge, S.P. (2005). Eye Movements and the Modulation of Parafoveal Processing By Foveal Processing Difficulty: A Re-examination. *Psychonomic Bulletin & Review, 12,* 891-896.

Yang, S.-N. (2006). An oculomotor-based model of eye movements in reading: The competition/interaction model. *Cognitive Systems Research, 7,* 56-69.

Yang, S.-N., & McConkie, G.W. (2001). Eye movements during reading: A theory of saccade initiation time. *Vision Research, 41,* 3567-85.

Appendix

Experimental sentences. The critical word is shown in italics. Sentences a, b, and c refer to the frequent and orthographically familiar, infrequent and orthographically familiar and infrequent and orthographically unfamiliar conditions respectively.

1a  He loved to visit the local *town* near to where his grandparents lived.

1b  He loved to visit the local *cove* near to where he learnt to swim.

1c  He loved to visit the local *quay* near to his father's fish shop.

2a  He thought the awful *party* was a waste of time.

2b  He thought the awful *lager* was really not good enough.

2c  He thought the awful *vinyl* was a lot worse than the old carpet.

3a  He tried to lift the heavy *door* onto its hinges but he needed more help.

3b  He tried to lift the heavy *gong* onto the table without it making a noise.

3c  He tried to lift the heavy *tusk* onto the truck but it was just too awkward.

4a  She thought the dusty *glass* might be very valuable.

4b  She thought the dusty *cello* might need re-tuning.

4c  She thought the dusty *yucca* might need watering.

5a  He saw the famous *place* in the city where the pop star was born.

5b  He saw the famous *adder* in the reptile section of the zoo.

5c  He saw the famous *crypt* in the old church.

6a  The photograph showed the young *child* sitting on top of the climbing frame.

6b  The photograph showed the young *heron* sitting on its nest.

6c   The photograph showed the young *koala* sitting in the eucalyptus tree.

7a   She liked the basic *home* because it was practical and well designed.

7b   She liked the basic *loom* because she could easily weave beautiful fabrics.

7c   She liked the basic *yoga* because it provided good but gentle exercise.

8a   She admired the unique *city* before she found out about the drug problem.

8b   She admired the unique *dart* before she aimed it at the dartboard.

8c   She admired the unique *oboe* before she began to play her favourite composition.

9a   Eventually the strong *wife* managed to move the concrete slab.

9b   Eventually the strong *lout* managed to push his way to the bar.

9c   Eventually the strong *oxen* managed to finish ploughing the field.

10a   She laughed at the funny *sound* that was coming from the radiator.

10b   She laughed at the funny *hound* that was bounding across the field.

10c   She laughed at the funny *scowl* that the child was making.

11a   She knew that the modern *chair* was perfect for her bedroom.

11b   She knew that the modern *grate* was of very poor quality.

11c   She knew that the modern *kiosk* was not suitable for the business.

12a   She looked at the awful *fire* which was spreading across the forest.

12b   She looked at the awful *lice* which were causing so many problems.

12c   She looked at the awful *acne* which she desperately wanted to be better.

13a   He used the decent *table* for the dinner party at his house.

13b   He used the decent *plank* for the building work on the garage.

13c   He used the decent *opium* for the very last time.

14a   He knew that the cheap *paper* might not be acceptable for the office.

14b  He knew that the cheap *crate* might not be strong enough to hold the bottles.

14c  He knew that the cheap *khaki* might not be quite the right colour.

15a  She chose the normal *bank* for the new business account.

15b  She chose the normal *sari* for her daughter to wear for the visit.

15c  She chose the normal *kiln* for the pots to be fired in.

16a  She liked the pretty *hair* that her sister had always had.

16b  She liked the pretty *glen* that she saw with her family in Scotland.

16c  She liked the pretty *loch* that was so quiet and peaceful.

17a  He gave the spare *money* to his brother after school.

17b  He gave the spare *cress* to his mother for the recipe.

17c  He gave the spare *myrrh* to the chemist at the university.

18a  She wanted the clean *room* ready for when her parents visited.

18b  She wanted the clean *lint* ready for the first aid session.

18c  She wanted the clean *toga* ready for the big party at the start of term.

19a  She liked the clever *idea* despite the expensive cost of the project.

19b  She liked the clever *dame* despite the annoying high pitched voice.

19c  She liked the clever *guru* despite his rather extreme beliefs.

20a  She worried about the major *case* that she had been asked to work on.

20b  She worried about the major *ford* that she would have to cross tomorrow.

20c  She worried about the major *feud* that had arisen within the family.

21a  She hated the awful *view* which she looked out on from her window.

21b  She hated the awful *lisp* which she had had since she was a child.

21c  She hated the awful *levy* which the society had decided to charge.

22a  He needed some normal *water* for the special mixture.

22b  He needed some normal *mince* for the special meal that he was cooking.

22c  He needed some normal *gauze* for the machine he was constructing.

23a  He disliked the boring *woman* despite her admirable achievements.

23b  He disliked the boring *chime* despite his love of clocks.

23c  He disliked the boring *polka* despite the fact that he was winning.

24a  The doctor looked at the unique *blood* under the microscope.

24b  The doctor looked at the unique *finch* under the old tree.

24c  The doctor looked at the unique *algae* under the water.

25a  She hated the yellow *house* which belonged to the dentist.

25b  She hated the yellow *froth* which clung to the edge of the bowl.

25c  She hated the yellow *fudge* which her grandmother gave her to eat.

26a  She admired the great *work* that the charity had done for the homeless.

26b  She admired the great *lark* that sang so beautifully in the tree.

26c  She admired the great *judo* that she saw the children doing in the arena.

27a  He planned the entire *story* before he spoke to the publisher.

27b  He planned the entire *prank* before he told his friends what to do.

27c  He planned the entire *haiku* before he wrote it for his girlfriend.

28a  He was impressed by the great *game* that everyone wanted to play.

28b  He was impressed by the great *solo* that the singer had performed.

28c  He was impressed by the great *hoax* that they had managed to pull off.

29a  He examined the clean *stage* before the performance.

29b  He examined the clean *quill* before he began to write.

29c  He examined the clean *anvil* before the smith used it to shape the metal.

30a  He wanted the quiet *night* to last forever because it was so peaceful.

30b  He wanted the quiet *tramp* to move away from the shop entrance.

30c  He wanted the quiet *hyena* to come back so that he could see him again.

31a  He thought the simple *light* was perfect for the office.

31b  He thought the simple *plait* was very appropriate on the little girl.

31c  He thought the simple *humus* was a bit boring for the special sandwiches.

32a  He cleaned the dirty *food* before giving it to the animals.

32b  He cleaned the dirty *leek* before he chopped it up for the stew.

32c  He cleaned the dirty *ruby* before he took it to the jewellers.

33a  She knew that the modern *road* would be safest in the winter conditions.

33b  She knew that the modern *cork* would keep the wine in good condition.

33c  She knew that the modern *tyre* would last a long time.

34a  She found the small *book* under the bed in her room.

34b  She found the small *dice* under the board game and cards.

34c  She found the small *kiwi* under the oranges and apples in the fridge.

35a  He watched the quiet *class* whilst they worked on the project.

35b  He watched the quiet *panda* whilst he hid amongst the bamboo bushes.

35c  He watched the quiet *waltz* whilst he waited at the dance hall.

36a  He knew that the single *girl* would not want his phone number.

36b  He knew that the single *pear*  would be good for the fruit crumble.

36c  He knew that the single *buoy* would be positioned near the stranded ship.

37a  She wanted a decent *world* for the poor people in the developing countries.

37b  She wanted a decent *puree* for the fantastic meal she was trying to prepare.

37c  She wanted a decent *dowry* for herself and her husband after the marriage.

38a  She used the fresh *fish* for the family dinner.

38b  She used the fresh *lard* for frying the meatballs.

38c  She used the fresh *suet* for making the dumplings.

39a  He examined the small *group* using a basic written test for each child.

39b  He examined the small *louse* using a special microscope.

39c  He examined the small *ulcer* using key-hole surgery.

Footnotes

[1] Note that type and token frequency are often confounded such that the term "orthographic regularity" is used rather than "orthographic familiarity" (White & Liversedge, 2006b). The term "orthographic regularity" can also be used to infer the extent to which letter sequences follow orthographic rules. However as the present paper focuses only on orthographic familiarity, this terminology is not used.

[2] Previous studies that have manipulated the number of higher frequency neighbours during sentence reading have shown either no effects (Sears, Campbell, & Luper, 2006) or late effects (Perea & Pollatsek, 1998, but see Pollatsek, Perea, & Binder, 1999, Sears et al., 2006). It is therefore unlikely that differences in the number of higher frequency neighbours could have produced the differences between the frequent and infrequent word conditions reported here. Nevertheless, any effects of the number of higher frequency neighbours would reflect a lexical level of processing, consistent with the reported conclusions related to effects of word frequency.

[3] The effects of linguistic variables on word skipping are often quite small. Therefore in order to increase the power of the analyses, each participant saw all three versions of each experimental item. Consequently the repeated sentence beginnings were specifically designed to be very bland and counterbalancing procedures ensured that the repeated items were spaced throughout the stimuli lists. The order in which the participants saw the different conditions for each item was also counterbalanced across the

three lists. After the experiment participants were asked if they noticed any kind of

repetition between the sentences and none of them did.

[4] The measures were also calculated separately for participants for whom the left

and right eyes were recorded, the two groups showed the same patterns of results.

[5] There is an issue about whether within-items or between-items designs should be

adopted when the critical words are different across conditions. A between-items design is

used when there are no controls for matching other variables across the conditions. In

contrast, when an experiment has a within-item design, the different items are matched

across the conditions, for example, by using the same sentence frame and matching for

other variables such as length of the critical words. In the present study, the overall pattern

of results was similar when the items analyses were undertaken with between-items tests.

For example, for orthographically familiar words, reading times on the critical word were

significantly longer for infrequent compared to frequent words for first fixation durations,

$t_2(76) = 5.55$, $p < 0.001$, single fixation durations, $t_2(76) = 6.14$, $p < 0.001$, gaze

durations, $t_2(76) = 7.3$, $p < 0.001$, and total time, $t_2(76) = 6.24$, $p < 0.001$. Other analyses,

such as the main effects of condition on word skipping, $F_2(2,114) = 2.23$, $MSe = 0.014$, $p$

$= 0.112$, partial $\eta^2 = 0.038$, and fixation durations prior to fixating the critical word

launched from word n-1, $F_2(2,114) = 2.4$, $MSe = 227$, $p = 0.095$, partial $\eta^2 = 0.04$, did not

reach significance for the between items analyses, though the same patterns of effects as

for the within-items tests clearly hold. Overall, the between-items analyses have larger $p$

values because of the less powerful design, that is, because between-items analyses do not

take account of the controls within the repeated-measures design. These analyses highlight

the importance of adopting careful matching procedures in studies of word recognition in natural reading.

[6] In an additional analysis, prior fixations were divided into those with near ($M = 3$, $SD = 1.4$) and far ($M = 6.7$, $SD = 3$) launch sites prior to fixating the critical word for each participant. There was no difference between the frequent and infrequent conditions for either near (Frequent: $M = 237$, $SD = 58$; Infrequent: $M = 238$, $SD = 64$) or far (Frequent: $M = 231$, $SD = 64$; Infrequent: $M = 231$, $SD = 61$) launch sites. Similar to the data for saccades launched from word n-1, prior fixation durations were numerically longer prior to orthographically unfamiliar (Near: $M = 242$, $SD = 65$; Far: $M = 238$, $SD = 63$) than familiar (Near: $M = 238$, $SD = 64$; Far: $M = 231$, $SD = 61$) words but none of the effects were statistically reliable. There was also no significant effect of condition on prior fixation durations launched three or less characters from the critical word ($F$s $< 1.43$, $p$s $> .24$).

[7] A 2 (skip or fixate critical word) X 3 (condition) ANOVA was undertaken to assess whether skipping the critical word influenced prior fixation durations. There was no effect of whether the critical word was fixated ($M = 240$) or skipped ($M = 239$) on the duration of the fixation prior to the critical word launched from word n-1, $F_1 < 1$, $F_2 (1,37) = 1.34$, $p = 0.254$, partial $\eta^2 = 0.035$. There was no effect of condition, $F_1 (2,52) = 1.43$, $MSe = 479$, $p = 0.248$, partial $\eta^2 = 0.052$, $F_2 < 1$, or an interaction, $F_1 (2,52) = 2.48$, $MSe = 577$, $p = 0.094$, partial $\eta^2 = 0.087$, $F_2 (2,74) = 1.9$, $MSe = 724$, $p = 0.157$, partial $\eta^2 = 0.049$. The hint of an interaction is likely to be due to the significant parafoveal-on-foveal

effect of orthographic familiarity for cases in which the critical word is fixated (as detailed in the Results). See Kliegl and Engbert (2005) for further discussion of this issue.

[8] Note that there were no differences in launch site between the three conditions prior to skipping or fixating the critical word, both for saccades launched from word n-1 ($F$s < 1) and from three or fewer characters away ($F$s < 1.3). In an additional analysis, fixations were divided into those with near ($M = 2.5$, $SD = 1.1$) and far ($M = 6.3$, $SD = 2.9$) launch sites for each participant, prior to the first pass of the critical word. There was no effect of condition on skipping probabilities for far launch sites ($F$s < 1.9, $p$s > 0.17) but there was an effect for near launch sites, $F_1 (2,58) = 3.13$, $MSe = 0.004$, $p = 0.05$, partial $\eta^2 = 0.097$, $F_2 (2,76) = 3.74$, $MSe = 0.005$, $p < 0.05$, partial $\eta^2 = 0.09$. Saccades launched from near launch sites were numerically more likely to skip the frequent (0.4) than the equally orthographically familiar infrequent words (0.35), though the effect was not significant, $t_1 (29) = 1.61$, $p = 0.118$, $t_2 (38) = 1.73$, $p = 0.091$. There was no significant difference in skipping probability between the orthographically familiar (0.35) and unfamiliar (0.32) infrequent words ($t$s < 1).

[9] The analysis of the probability of making a regression in to the critical word was based on 28 participants in the analyses across participants, and 38 items in the analyses across items for the main 2 X 3 ANOVA and for the three way ANOVA for cases in which the critical word was skipped on first pass.

Table 1. Non-position Specific (NPS) and Position Specific (PS) Summed Token N-gram Frequency Counts for Each of the

Conditions. Standard Deviations Shown in Parentheses. Difference in Mean Frequency Counts for the Orthographically Familiar

Conditions and the Infrequent Conditions.

| Condition | Frequent, orthographically familiar | | Infrequent, orthographically familiar | | *Difference: Orthographically familiar words* | | Infrequent, orthographically unfamiliar | | *Difference: Infrequent words* | |
|---|---|---|---|---|---|---|---|---|---|---|
| N-gram | NPS | PS | NPS | PS | *NPS* | *PS* | NPS | PS | *NPS* | *PS* |
| Trigram | 5654 | 2905 | 7369 | 2668 | *1715* | *-237* | 541 | 231 | *-6828\*\*\** | *-2437\*\*\** |
| | (3842) | (2135) | (6513) | (3173) | | | (573) | (310) | | |
| Bigram | 65315 | 21798 | 84564 | 23246 | *19249\*\** | *1448* | 33422 | 10109 | *-51142\*\*\** | *-13137\*\*\** |
| | (30237) | (11593) | (41470) | (9713) | | | (22788) | (8767) | | |
| Monogram | 1154817 | 250301 | 1181874 | 242741 | *27057* | *-7560* | 989524 | 205336 | *-192350\*\** | *-37405\** |
| | (227950) | (66055) | (245519) | (49672) | | | (247410) | (68417) | | |

Note: \* = $p < .05$, \*\* = $p < .01$, \*\*\* = $p < .001$.

*Table 2.* First Fixation Durations, Single Fixation Durations and Gaze Durations on Word n-1. Mean Fixation Duration Prior to Fixating the Critical Word for All of the Data, for Only Saccades Launched from Word n-1, and for Only Saccades Launched from Word n-1 Except From the Final Three Characters of the Word. Standard Deviations Shown in Parentheses. The Size of the Frequency and Orthographic Familiarity Effects are Shown in Italics.

| Reading time measure word n-1 | Frequent, orthographically familiar | Infrequent, orthographically familiar | *Frequency effect* | Infrequent, orthographically unfamiliar | *Orthographic familiarity effect* |
|---|---|---|---|---|---|
| First fixation duration | 243 (60) | 242 (59) | *-1* | 244 (63) | *2* |
| Single fixation duration | 245 (59) | 246 (58) | *1* | 248 (63) | *2* |
| Gaze duration | 263 (82) | 264 (78) | *1* | 265 (85) | *1* |
| Prior fixation duration (all) | 233 (62) | 233 (64) | *0* | 239 (65) | *6* |
| Prior fixation duration (n-1) | 238 (60) | 238 (60) | *0* | 244 (63) | *6* |
| Prior fixation duration (n-1, except final 3 characters) | 240 (58) | 239 (55) | *-1* | 246 (60) | *7* |

*Table 3*. Mean Reading Times on the Critical Word for Each Condition. Probability of Refixating the Critical Word on First Pass. Mean Duration of the Spillover Fixation. Standard Deviations Shown in Parentheses. The Size of the Frequency and Orthographic Familiarity Effects are Shown in Italics.

| Reading measure | Frequent, orthographically familiar | Infrequent, orthographically familiar | *Frequency effect* | Infrequent, orthographically unfamiliar | *Orthographic familiarity effect* |
|---|---|---|---|---|---|
| First fixation duration | 253 (77) | 280 (89) | *27* | 286 (93) | *6* |
| Single fixation duration | 255 (78) | 284 (88) | *29* | 294 (92) | *10* |
| Gaze duration | 265 (88) | 309 (117) | *44* | 324 (135) | *15* |
| Total time | 289 (123) | 356 (168) | *67* | 365 (181) | *9* |
| Refixation | 0.06 | 0.11 | *0.05* | 0.15 | *0.04* |
| Spillover fixation duration | 234 (66) | 245 (83) | *11* | 241 (77) | *-4* |

*Table 4.* Probability of Skipping the Critical Word for All of the Data, for Only Saccades Launched from Word n-1, and for Only

Saccades Launched from Three or Less Characters from the Critical Word. Mean Landing Position, Launch Site and Saccade

Length for the Critical Word. Standard Deviations Shown in Parentheses. Probability of First Pass Regressions Out of the Critical

Word, and in to the Critical Word for when the Critical Word was Fixated (Fix n) and Skipped (Skip n) on First Pass. The Size of

the Frequency and Orthographic Familiarity Effects are Shown in Italics.

| Reading measure | Frequent, orthographically familiar | Infrequent, orthographically familiar | *Frequency effect* | Infrequent, orthographically unfamiliar | *Orthographic familiarity effect* |
|---|---|---|---|---|---|
| Skip (all data) | 0.23 | 0.2 | *- 0.03* | 0.19 | *- 0.01* |
| Skip (launch n-1) | 0.26 | 0.22 | *- 0.04* | 0.2 | *- 0.02* |
| Skip (launch ≤ 3 characters) | 0.41 | 0.34 | *- 0.07* | 0.3 | *- 0.04* |
| Landing position | 3.0 (1.3) | 3.0 (1.3) | *0* | 3.0 (1.3) | *0* |
| Launch position | 4.0 (1.6) | 3.9 (1.6) | *- 0.1* | 3.9 (1.7) | *0* |
| Saccade length | 7.1 (1.5) | 7.0 (1.6) | *- 0.1* | 6.9 (1.5) | *- 0.1* |
| First pass regression out | 0.08 | 0.09 | *0.01* | 0.09 | *0* |
| First pass regression in (Fix n) | 0.09 | 0.12 | *0.03* | 0.11 | *-0.01* |
| First pass regression in (Skip n) | 0.26 | 0.45 | *0.19* | 0.52 | *0.07* |

Figure Caption

Figure 1. The distribution of single fixation durations on the critical word for each of the three experimental conditions (35ms bins).

Figure 1.



**Single fixation duration (ms)**

Legend:
- Frequent, orthographically familiar
- Infrequent, orthographically familiar
- Infrequent, orthographically unfamiliar