Article

Parafoveal preview effects in reading unspaced text

Drieghe, Denis, Fitzsimmons, Gemma and Liversedge, Simon Paul

Available at http://clok.uclan.ac.uk/22324/

It is advisable to refer to the publisher's version if you intend to cite from the work.
http://dx.doi.org/10.1037/xhp0000441

For more information about UCLan's research in this area go to http://www.uclan.ac.uk/researchgroups/ and search for <name of research Group>.

For information about Research generally at UCLan please go to http://www.uclan.ac.uk/research/

CLoK
Central Lancashire online Knowledge
www.clok.uclan.ac.uk

CLoK
University of Central Lancashire's Institutional Repository

Parafoveal Preview Effects in Reading Unspaced Text

Denis Drieghe, Gemma Fitzsimmons, & Simon P. Liversedge

University of Southampton

Contact details:

Denis Drieghe

School of Psychology

University of Southampton

Southampton, SO17 1BJ

UK

d.drieghe@soton.ac.uk

+44 238095 2585

Abstract

In English reading, eye guidance relies heavily on the spaces between words for demarcating word boundaries. In an eye tracking experiment, we examined the impact of removing spaces on parafoveal processing. Using the gaze-contingent boundary paradigm (Rayner, 1975), a high or low frequency pre-boundary word was followed by a post-boundary preview presented either normally (i.e. identical to the post-boundary word), or with letters replaced creating an orthographically illegal preview. The spaces between words were either retained or removed. Results replicate previous findings of increased reading times during unspaced reading (Rayner, Fischer & Pollatsek, 1998) and indicate rather limited evidence for more distributed processing: Observations of processing of the previous word (spill-over effects) or processing of the next word (parafoveal-on-foveal effects) influencing fixation durations on the currently fixated word were limited. Spill-over effects were only observed in the unspaced layout when the post-boundary preview was correct, presumably because the orthographically illegal, incorrect preview was visually salient enough to allow for relatively easy word segmentation and therefore more focused processing of the pre-boundary word. As such, results points towards a system that prefers narrowly focused processing of a single word, at least when means for easy word segmentation are available.

Statement of the Public Significance of the Work.

In the influential gaze contingent boundary paradigm (Rayner, 1975), before the eyes cross an invisible boundary during reading, a preview is presented at the location of the target word that can either be identical to the target word or manipulated to be related to a certain extent with the target word (e.g. number of shared letters). After the eyes cross the boundary, a display change is carried out replacing the preview with the target word. Using this paradigm we examined the impact of removing spaces between the words. Results replicate previous findings of increased reading times during unspaced reading and indicate rather limited evidence for distributed processing when means for easy word segmentation were available (either spaces or an orthographically illegal string of letters in the preview).

Reading is a complex, cognitive task that has to accommodate a limitation in our visual system in that the fine-grained visual acuity necessary to identify letters is mostly restricted to a limited area of the retina, the *fovea*, which receives input from only the central 2 degrees of the visual field (for reviews, see Rayner, 1998; Rayner, 2009). As a consequence, reading typically requires a multitude of fixations and saccades to allow sampling of visual information at different locations across the text. Nevertheless, a large body of evidence indicates that besides the currently fixated word, readers do routinely pick up information from the upcoming word as well (for a review, see Schotter, Angele, & Rayner, 2012). The upcoming word will typically be in an area called the *parafovea* that extends from the fovea to 5 degrees of the centre of the visual field. The fact that readers do pick up useful information from the parafoveal word and use this information during the subsequent fixation on that word is apparent from research using the *gaze-contingent boundary* paradigm (Rayner, 1975; Schotter et al., 2012). In this influential paradigm, an invisible boundary is placed before the target word. Before the eyes cross the invisible boundary, a preview is presented at the location of the target word that can either be identical to the target word or manipulated to be related to a certain extent with the target word (e.g. number of shared letters). After the eyes cross the boundary, a display change is carried out replacing the preview with the target word. Because this display change happens during a saccade when the eyes are functionally blind due to saccadic suppression (Matin, 1974), participants are typically unaware of this manipulation. The *parafoveal preview benefit* consists of the reduction of fixation times on the target word observed when the preview was the correct target word compared to when the preview was different from the target word. It is this preview benefit that is taken as evidence that useful information from the parafoveal word is already being extracted before the

word is directly fixated (for a recent meta-analysis examining different preview manipulations, see Vasilev & Angele, 2017).

A long outstanding question in the field of eye movements during reading (e.g. Starr & Rayner, 2001) concerns whether lexical processing of words happens in a serial fashion or whether multiple words are processed simultaneously. Serial models, such the E-Z Reader model (Reichle, Pollatsek, Fisher & Rayner, 1998; Reichle, Warren & McConnell, 2009), advocate that parafoveal processing of the upcoming word does not begin before lexical processing of the currently fixated word has been concluded. Parallel models, such as the SWIFT model (Engbert, Nuthmann, Richter, & Kliegl, 2005; Schad & Engbert, 2012), posit that all words within the perceptual span are to varying extents simultaneously lexically processed. An issue that is very likely to have an impact on the extent to which words are processed (or indeed can be processed) serially or in parallel is the ability of the visual and linguistic processing system to perform fast and efficient word segmentation, and as such to allow for narrow, focused foveal and parafoveal processing. English, as in most alphabetic languages, is spaced and spaces will afford such straightforward word segmentation (see below). However, many writing systems are unspaced, such as Chinese, and do not have overt word boundary demarcation in their written form. In this study, we will examine the role of spacing in allowing selective focusing of foveal processing on the currently fixated word and parafoveal processing on the upcoming word.

Spaces provide unambiguous cues to word boundaries (with notable exceptions such as spaced compounds) and provide low spatial frequency information about the length and location not just of the currently fixated word but also for the upcoming words. Using the moving window paradigm in which the reader is allowed correct information only about the currently fixated word and a specific area around it,

McConkie and Rayner (1975) were able to determine that reading fluency was negatively influenced when masks did not preserved spacing information up to 15 characters to the right of fixation. This negative impact on reading fluency indicates that readers under normal circumstances do extract spacing cues and utilise word length information even at such a distance from the current fixation location (for a review, see Rayner, 2014).

Another way in which researchers have examined the importance of spacing has been to remove the spaces from text and examine eye movement behaviour during unspaced reading. Removing spaces in English results in an increase in reading times by 30-50% and it also increases average fixation durations, increases the number of fixations, increases the number of regressions, and reduces the average saccade length (Epelboim, Booth, Askenazy, Talghani & Steinman, 1997; Epelboim, Booth & Steinman, 1994; Malt & Seaman, 1978; McGowan, White, Jordan, & Paterson, 2014; Perea & Acha, 2009; Perea, Tejero & Winskel, 2015; Pollatsek & Rayner, 1982; Rayner, Fischer & Pollatsek, 1998; Rayner & Pollatsek, 1996; Sheridan, Rayner, & Reingold, 2013; Sheridan, Reichle & Reingold, 2016; Spragins, Lefton & Fisher, 1976; Veldre, Drieghe, & Andrews, 2017; Yang & McConkie, 2001).

A number of factors have been suggested in the literature as potential causes for the beneficial effects of spacing in languages that are normally spaced (e.g. Pollatsek & Rayner, 1982; Rayner, Fisher & Pollatsek, 1998; Sheridan, Rayner & Reingold, 2013). First of all, spaces in all likelihood reduce visual crowding and lateral masking for the outer letters of words. Studies that have increased spacing have shown small but consistent benefits on reading speed (e.g. Drieghe, Brysbaert & Desmet, 2005; Slattery & Rayner, 2013). Second, spacing aids in guiding eye movements by providing word boundaries and word length information and removing

them disrupts oculomotor planning. To be more specific, in English, saccades typically tend to land around an area called the *preferred viewing location* (Rayner, 1979), which is slightly to the left of the centre of a word. When reading unspaced text, readers tend to land comparatively closer to the beginning of words (e.g. Perea & Acha, 2009; Sheridan, Rayner, & Reingold, 2013) and this has been taken as an indication that when spaces are available they are used in guiding the eyes towards the position that is ideal for lexical processing. It has also been suggested that spacing helps lexical word identification by providing information on where a word begins and where it ends. Rayner et al. (1998) reported a larger frequency effect (a high-frequency word typically receives shorter fixation times than a low-frequency word, e.g. Inhoff & Rayner, 1986; Rayner & Duffy, 1986) for unspaced reading compared to spaced reading in gaze duration. These findings, both the increased frequency effect in unspaced compared to spaced reading, and the time course of the effect in that it appears in gaze duration and not earlier, has been replicated multiple times (e.g. Perea & Acha, 2009; Sheridan, Rayner & Reingold, 2013).

Additionally, using survival analyses, Sheridan et al. (2013) were able to show a later time course of word frequency effects in unspaced versus spaced reading. They observed the earliest discernible influence of word frequency in spaced text 112 ms from the start of fixation, with no discernible difference until 152 ms into the fixation when the text was unspaced. Finally, in an individual differences study, Veldre, Drieghe and Andrews (2017) observed that the negative impact of the unspaced format on reading behaviour was less pronounced for participants with good spelling ability, again linking the effects of spacing to lexical processing, more specifically to the quality of the lexical representations as indexed by spelling ability.

At this point it may be useful to consider how readers of unspaced languages deal with the issue of word segmentation. Several studies have shown that the size of preview effects tends to be larger for reading in Chinese compared to alphabetic languages (e.g. Yan, Richter, Shu & Kliegl, 2009; Yen, Radach, Tzeng, Hung & Tsai, 2009). This finding was recently corroborated in a meta-analysis by Vasilev and Angele (2017) but with the caveat that the evidence for this larger preview effect was mostly restricted to observations in gaze duration. It has been suggested that the reason for this comparatively larger preview effect in Chinese is the unspaced nature of Chinese script. Theoretically, one could imagine that the processing of Chinese could be entirely character based such that the stream of Chinese characters would not need to be segmented into words. However, a large body of evidence clearly demonstrates the psychological reality of words in Chinese (for a review see Zang, Liversedge, Bai & Yan, 2011) and as such demonstrates the necessity of word segmentation during Chinese reading. This necessity could have a profound impact on the extent to which foveal words and parafoveal words are lexically processed serially or in parallel. That is, the lack of low-level visual boundaries may result in readers processing multiple characters in parallel – and across word boundaries - to a larger extent in order to allow for word segmentation compared to spaced languages where word segmentation can be straightforward and performed quickly. Indeed, Yen et al. (2009) suggested that for an unspaced script such as Chinese, parallel character processing in the parafovea is a sensible starting point for modelling eye movement behaviour.

Cui et al. (2014) examined the issue of the impact of the unspaced format on parafoveal processing by implementing a boundary change experiment in Chinese, both in the traditional, unspaced format and in a format in which spaces were inserted

in between the words. They observed that for one character words the insertion of spaces before and after the word did result in an increased preview effect - above what could be expected solely on the basis of reduced visual crowding - and proposed that the spacing allowed for more narrowly focused and more efficient parafoveal processing.

In the current experiment, we will examine the influence of spacing on parafoveal processing by carrying out the mirror image of the Cui et al. (2014) study. Whereas Cui et al. inserted spaces into Chinese text, an unspaced language, the current study will implement a boundary change experiment in English, a spaced language, comparing preview effects in the familiar, spaced layout with an unspaced layout. The only study we are aware of that carried out a similar manipulation was reported by Sheridan, Reichle and Reingold (2016). They implemented a gaze-contingent boundary paradigm for reading text in which either the spaces were preserved or in which the spaces were replaced by random numbers. Sheridan et al. reported an interaction between the preview of the target word and the layout of the text in the first fixation duration on the target word such that the preview effect was reduced in unsegmented compared to spaced text. This interaction was marginally significant in gaze duration and not significant in single fixation duration on the target word.

The current study will expand the experiment from Sheridan et al. by removing the spaces altogether instead of replacing them by random numbers. As such, analogous to unspaced languages, any demarcation signals between words will be removed. Note that the differences between the Sheridan et al. and the current study are non-trivial given the differences in how the unspaced layout was implemented. Veldre, Drieghe and Andrews (2017) directly compared spaced reading

with three unsegmented text conditions that either preserved or eliminated word boundary information: a condition in which the spaces were replaced by numbers, a condition in which the spaces were removed but words were presented alternating in either uppercase or lowercase for the entire word, and an unspaced condition in which the spaces were removed and no additional demarcation information was added. The condition in which the spaces were replaced by numbers was identical to the manipulation used by Sheridan et al. (2016) and resulted still in considerably faster reading times than the unspaced condition. Whereas numbers are in all likelihood very effective in removing word length information well into the parafovea, when the eyes are very close to them, they might still provide some word demarcation cues. We therefore opted for the more straightforward removal of spaces between the words.

We also manipulated the pre-boundary word such that it was either a high- or a low-frequency word. Henderson and Ferreira (1990) were the first to show a reduced preview effect when the pre-boundary word was low-frequency compared to when it was high-frequency. Models such as E-Z Reader (Reichle et al., 1998) explain this finding by assuming that the processing of the pre-boundary word takes comparatively longer when it is low-frequency, causing the time window in which parafoveal processing can occur to be reduced (this time window is between the arrival of attention on the post-boundary word after the pre-boundary word has been recognised, and the arrival of the eyes on the post-boundary word). Most parallel models (e.g. SWIFT, Engbert, Nuthmann, Richter, & Kliegl, 2005) explain this finding by assuming that the processing of a low-frequency word will take up more resources, reducing the speed with which other words are processed. An effect of the frequency of a word on the fixation time of the next word has been called a *spill-over effect* and our current design will allow us to examine these effects in the context of

spaced versus unspaced reading. Summarising, the current study will feature a boundary change that occurs during reading with the frequency of the pre-boundary word being either high or low, with the layout being either spaced or unspaced, and with the preview of the post-boundary word being either identical to the post-boundary word or a non-word (for an example of the stimuli, see Table 1).

INSERT TABLE 1 ABOUT HERE

Focusing on the eye movement behaviour on the pre- and post-boundary word, in this experiment we expect to replicate the classic findings associated with reading unspaced text compared to spaced text: Longer fixation times, initial landing positions closer to the word beginning, reduced skipping rate, a higher number of fixations when fixated and an increased frequency effect (see literature above). For the impact of spacing on serial versus parallel lexical processing, we will focus on two specific phenomena that can be expected to increase as simultaneous lexical processing of multiple words occurs to a higher extent than during traditional, spaced reading. The first is the processing of the previous word influencing the fixation times on the currently fixated word (i.e. spill-over effects) and the second one is the processing of the upcoming word influencing the fixation times of the currently fixated word (so-called *parafoveal-on-foveal effects*, the existence of which during spaced reading experiments has been controversial, for a review see Drieghe, 2011).

With regard to the preview manipulations, two hypotheses can be created. First of all, due to a lack of low-level word boundary demarcation in unspaced text, lexical processing is forced to be - to a greater degree than during normal, spaced reading – distributed across multiple words, leading to comparatively more parafoveal-on-foveal and spill-over effects. Note that a distinction can be made

between predictions indicative of more parafoveal processing versus less narrow processing. In the case of more parafoveal processing we would expect a larger preview effect (i.e. difference between the fixation times on the post-boundary word when the preview was correct versus when the preview was the non-word), in the case of less narrow processing we would expect more parafoveal-on-foveal and spill-over effects. As such, this distinction can also be thought of as a difference between depth of processing and extent of processing. The first hypothesis predicts less narrow processing. The alternative hypothesis can also be proposed, whereby, compared to spaced reading, lexical processing of the foveal and parafoveal word might happen in a (more) serial fashion. That is, the system quickly establishes that little can be gained from parafoveal processing (indeed, in the current experiment this will occasionally lead to the processing of non-words), leading to a strategy whereby the system conservatively tries to put a word boundary as early as possible in the upcoming stream of letters, and will primarily focus its resources on foveal processing. This would lead to no, or very limited, spill-over effects and parafoveal-on-foveal effects.

METHOD

Participants

Twenty-four native English speakers with normal or corrected-to-normal vision from the University of Southampton participated for £4.50 or course credit.

Apparatus

Eye movements were recorded with an SR Research Eyelink 1000 system. Viewing was binocular, but eye movements were recorded from the right eye only.

Sentences were displayed on a single line with a maximum length of 85 characters. The letters were presented in monospaced Courier font, were lowercase (except when capitals were appropriate) and were presented in black on a grey background. The display was about 73 cm from the participants' eye and at this distance 3 characters equalled 1° of visual angle.

Materials and design

The stimuli were made up of 60 sentence frames. A pre-boundary word was embedded in the neutral sentence. The pre-boundary word was 5 letters long and was either a high-frequency noun (4.8 Zipfian Log frequency in the SUBTLEX-UK Database; Van Heuven, Mandera, Keuleers, & Brysbaert, 2014) or a low-frequency noun (3.5 Zipfian Log frequency). The difference in frequency was significant ($t(59)$ = 16.03, $p < .001$) and the value of 3.5 Zipfian Log Frequency indicates that our low-frequency words were not very low frequency (A Zipfian value of 1-3 is typically considered low-frequency, a value of 4-7 is considered high-frequency). We selected these words to ensure that they were recognizable in the unfamiliar unspaced format. The sentences were presented either with normal spacing or with all spaces removed. Additionally, the preview of the post-boundary word before the boundary was crossed was either presented normally or was changed into a non-word. More specifically, the post-boundary word was always minimally 5 letters long, and in order to create the non-word the first 4 letters were replaced to produce a non-pronounceable non-word. As such the design was a 2 (Layout: spaced or unspaced) x 2 (Frequency of the pre-boundary word: High-Frequency or Low-Frequency) x 2 (Preview: Normal or Non-Word Preview) within-subjects design. Participants saw only one of the 8 possible versions of each sentence. These 60 sentences were mixed with 60 filler items (30

spaced, 30 unspaced) and were displayed in a pseudorandom order preceded by 16 practice sentences (8 spaced, 8 unspaced).

Procedure

Participants were first given a description of the experimental procedure and they were told they would be reading sentences on the monitor. They were instructed to read for comprehension and were also informed that they occasionally would be asked comprehension questions about the sentences. Participants were warned some of the sentences would be unspaced and that their task was to continue reading for comprehension to the best of their ability. The participant's head was stabilized using a head- and chinrest. The initial calibration took approximately 5 minutes. At the beginning of each trial the participant had to look at a fixation point on the left of the screen. When the tracker recorded a stable fixation on the fixation point, the sentence was displayed such that the fixation fell at the beginning of the sentences. When participants finished reading the sentence, they pressed a button on the response box to move on to the next trial. Comprehension questions were presented after 25% of the trials and the accuracy in answering them was 93%. The experiment lasted approximately 40 minutes.

RESULTS

Fixations shorter than 80ms that were within one character of a previous or subsequent fixation, were combined with that fixation. All other fixations shorter than 80ms were removed, as were those fixations whose durations exceeded 800ms. Trials were removed for the following reasons: a) Calibration error, b) The display change

was triggered during a fixation before the boundary, c) The display change was triggered by so-called hooking, in which a saccade crosses the boundary but hooks back landing before the boundary, and d) The display change occurred during a fixation located after the boundary relatively late into that fixation (more than 10 ms into the fixation; see Slattery, Angele & Rayner, 2011). Combined, 28.9% of trials were removed from the analyses[1]. Finally, for each measurement we also removed observations that were 2.5 standard deviations from the mean of each participant[2].

Analyses were conducted on the eye-movement measures both on the pre-boundary and the post-boundary word. These measures were *first fixation duration* (the duration of the first fixation on the target word, regardless of how many fixations the target word received), *single fixation duration* (the duration of the fixation on the target word when it received exactly one fixation), *gaze duration* (the sum of first-pass fixations on the target word), *go-past time* (the time between first fixating a target word and moving past it, which includes regressions launched from the target word), the *skipping rate* (the likelihood of the target word not receiving a fixation during first-pass), the *Number of Fixations* the target word receives during first-pass and for the pre-boundary word the *Initial Fixation Location* which is the location where the saccade onto the target word lands and is expressed in number of pixels to the right from the start of the target region.

The data were analyzed using Linear Mixed-effects Models (LMM) using the lme4 package (Version 1.1-12, Bates, Maechler, Bolker, & Walker, 2015) in R (Version 3.3.1; R Core Team 2016). Contrasts were specified as -.5/.5 and were used for the effects of Layout, Frequency and Preview such that the intercept corresponds to the grand mean and the fixed effects correspond to the main effect of the fixed factors. Fixation durations were log-transformed to increase normality of the data. For

the skipping data, logistic GLMMs were carried out given the binary nature of the dependent variable and for the count data (number of fixations) the data were modelled following a Poisson distribution. In addition, we entered subjects and items as crossed random effects. The random effects structure of the model consisted of slopes for all the fixed effects across subjects and items, including correlations (see Barr, Levy, Scheepers & Tilly, 2013), but was further trimmed down for those models that either did not converge or had correlations in the random structure equal to zero or one (which is a sign of overparametrisation)[3]. Absolute values of the t-value equal to or greater than 1.96 were interpreted as significant because for high degrees of freedom the *t* static in LMMs approximates the *z* statistic.

Pre-Boundary Word

Reading measures on the pre-boundary word are shown in Table 2, results of the LMM's are reported in Table 3.

INSERT TABLE 2 AND 3 ABOUT HERE

First Fixation Duration. There was a main effect of frequency such that the high frequency pre-boundary word received a first fixation that was on average 19ms shorter than the low-frequency pre-boundary word. A strong effect of layout was also observed in first fixation with fixation times 86ms shorter in the spaced compared to the unspaced layout. There was no effect of the preview manipulation and none of the interactions were significant.

Single Fixation Duration. In single fixation duration there was also an effect of frequency with a 24ms longer fixation on the low-frequency word compared to the high frequency word. A main effect of Layout was also observed. However, this main

effect was qualified by an interaction with Preview. This two-way interaction is depicted in Figure 1 and clearly shows how in the unspaced format there is a tendency for a shorter single fixation duration prior to the non-word preview compared to the normal preview. This was also clear from an analysis run separately on the spaced and unspaced dataset: The effect of preview was not significant in spaced layout (b = 0.04, SE = .0.03, t = 1.16) but was in the unspaced layout (b = -0.10, SE = 0.04, t = -2.40). None of the other interactions were significant.

INSERT FIGURE 1 ABOUT HERE

Gaze Duration. The patterns observed in gaze duration closely resemble the effects observed in first fixation duration: A significant effect of frequency such that a high-frequency word received gaze durations on average 66 ms shorter than the low-frequency-word and gaze durations were 170 ms shorter in spaced compared to unspaced reading. None of the interactions were significant which means that the numerically strong interaction in the raw data between frequency and layout, which indicated a stronger frequency effect in unspaced text (91 ms compared to 41 ms in spaced layout), was not significant in the LMM analysis of the log-transformed data. We will return to this surprising finding in the Discussion.

Go Past Times. A significant main effect of Layout was observed with longer fixation times in the unspaced compared to the spaced format and a significant main effect of Frequency with shorter fixation times on high-frequency words compared to low-frequency words. However, both effects were qualified by an interaction between Frequency and Layout. As can be seen from Figure 2, this interaction reflects a stronger frequency effect in the unspaced format. In the raw data, the frequency effect was on average 42ms in spaced layout and 199ms in unspaced layout.

INSERT FIGURE 2 ABOUT HERE

Skipping Rate. A high-frequency word was skipped 6% more often than a low-frequency word and skipping rates were 7% higher in spaced format compared to when the spaces were removed. There was no main effect of preview and none of the interactions were significant.

Number of Fixations. When the target word was fixated it received 0.22 more fixations in the unspaced layout compared to the spaced layout and a low-frequency word received 0.16 more fixations than a high-frequency word. There was no effect of preview and none of the interactions were significant.

Initial Landing Site. There was a main effect of Layout such that the eye landed 6 pixels earlier in the unspaced format compared to the spaced format. No other main effects or interactions were significant.

Post-Boundary Word

Reading measures on the post-boundary word are shown in Table 4, results of the LMM's are reported in Table 5.

INSERT TABLE 4 AND 5 ABOUT HERE

First Fixation Duration. A significant main effect of Preview and Layout is observed in the expected direction of longer fixations durations when the preview was a non-word compared to when the preview was presented normally and longer fixation durations in the unspaced compared to the spaced format. These two main effects were qualified by a three-way interaction between Layout, Frequency and Preview. This three-way interaction is displayed in Figure 3. The interaction is caused by a comparatively larger preview effect in those conditions where the pre-boundary

word was high frequency and the layout was unspaced. This is also apparent from the raw data where the preview effect for the unspaced high-frequency condition is 68ms, whereas it varies between 22-38ms for the other conditions.

INSERT FIGURE 3 ABOUT HERE

Single Fixation Duration. An almost identical pattern to that observed in first fixation durations emerges from the single fixation duration data. Main effects of Preview (longer times after a non-word preview compared to the normal preview) and Layout (longer fixations in unspaced compared to spaced reading) are qualified by a three-way interaction between Layout, Frequency, and Preview. This interaction is shown in Figure 4 and was again caused by a larger preview effect when the pre-boundary word was high-frequency in the unspaced layout compared to the preview effect in the other conditions. Again, this is also apparent from the raw data where the preview effect for the unspaced high-frequency conditions is 71ms whereas it varies between 37-45ms for the other conditions.

INSERT FIGURE 4 ABOUT HERE

Gaze Duration. There was a main effect of layout such that gaze durations were 129ms longer in unspaced format than spaced format and a main effect of preview such that the participants looked 44 ms longer at the post-boundary word when the preview was the non-word as compared to when it was the normal preview. There was no main effect of frequency and none of the interactions were significant.

Go Past Times. The patterns observed in Go Past Times closely resemble those observed in Gaze Duration. The post-boundary word received go-past times 397ms longer in the unspaced layout compared to the spaced layout and it received go-past

times 56ms longer when the preview was the non-word compared to when the preview was the normal preview. There was also a main effect of frequency such that longer go-past times were observed on the post-boundary word after a low-frequency pre-boundary word compared to after a high-frequency pre-boundary word. A marginally significant interaction indicates that these increased go-past times after a low-frequency word were mostly present in the unspaced layout, as is also strongly indicated by the raw fixation times. As these patterns were not present during the earlier fixation measures on the post-boundary word this indicates that the effect in go-past times is due to relatively later effects (more regressions and longer regressions) originating from the post-boundary target when the pre-boundary word was low-frequency. None of the other interactions were close to significant.

Skipping Rate. A main effect of spacing indicated that in the unspaced layout the post-boundary word was skipped 8 % less often than in the spaced layout. When the preview was correct, the post-boundary word was skipped more 7% often than when the preview was the non-word. The frequency of the pre-boundary word did not significantly influence the skipping rate of the post-boundary word and none of the interactions were significant.

Number of Fixations. When the post-boundary word was fixated it received 0.22 more fixations in the unspaced format compared to the spaced format. The main effects of frequency and preview were not significant, and neither were any of the interactions.

DISCUSSION

This experiment examined foveal and parafoveal processing during reading of English presented in either the familiar, spaced layout or with all the spaces between the words removed. The pre-boundary word was either a high- or a low-frequency word and the preview of the post-boundary word was either identical to the post-boundary word or letters were replaced creating a non-pronounceable non-word. The goal was to examine the impact of removing the easy and fast word demarcation offered by spaces on the ability to narrowly focus processing on the currently fixated word and parafoveal processing on the next. We assumed that less narrow focused processing would be associated with an increase of the processing of either the preceding (spill-over effects) or upcoming (parafoveal-on-foveal effects) word influencing the fixation times on the currently fixated word.

Starting with the observations on the pre-boundary word, a very similar picture emerged from the analysis of first fixation duration, gaze duration, the skipping rate and the number of fixations on the pre-boundary word. Compatible with findings reported in the literature (see Rayner, 2009 for a review), a high-frequency word received shorter first fixations and gaze durations, was skipped more often and when it was fixated received less fixations than a low-frequency word. The unspaced layout was associated with longer first fixations times and gaze durations, less skipping of the pre-boundary word and a higher number of fixations when the word was fixated. These findings replicate a large number of studies that examined unspaced reading (e.g. Rayner et al. 1998; Sheridan et al., 2016; Veldre, et al., 2016). Notably, on all these measurements there was no effect of the preview manipulation on the next word, in other words we did not find any parafoveal-on-foveal effects.

The go-past times on the pre-boundary word showed the same effects of preview and frequency, but in this measurement we observed the expected frequency

and layout interaction such that there was a larger frequency effect in unspaced compared to spaced reading (e.g. Rayner et al. 1998). What is inconsistent with the literature is that in previous studies, the first appearance of this interaction has been consistently reported in gaze duration (Perea & Acha, 2009; Rayner et al., 1998; Sheridan et al., 2013; Sheridan et al, 2016; Veldre et al. 2016), whereas this interaction was not close to significant for gaze durations in the current study. The absence of this interaction was even more surprising given the sizeable numerical interaction in the untransformed gaze durations (91ms frequency effect in unspaced compared to 41 ms in spaced reading). At this point it is important to note that the majority of previous studies reporting the frequency and layout interaction used ANOVA's to analyse the fixation durations and did not apply any transformation to the raw data. A re-analysis of our current data with an LMM on the untransformed data did show a significant interaction between frequency and layout (b = 45.61, SE = 21.62, t = 2.11), confirming that the interaction was removed by the transformation process, and that had we not carried out this procedure, our findings would have been consistent with previous studies.

Note that we are not cherry picking results here from the LMM's run either on the untransformed or transformed data, we are just establishing whether our experiment replicates previously reported findings, which it does but only when analysed the same way as in the previous studies. For our data, visual inspection clearly indicated an improvement in the extent to which our data followed a normal distribution, indicating that the LMM on the log transformed data is the most appropriate statistical model for analysing these data. However, we would also like to point out that if an effect is indeed exclusively in the right tail of the distribution (and as such removed by log-transformation), this does not necessarily imply that the effect

in the right tail was not psychologically relevant or interesting. We maintain our current choice of statistical model but would like to point the interested reader to a paper by Balota, Aschenbrenner, and Yap (2013) where concerns are voiced towards the standard procedure of transforming data prior to running LMM's. One could establish the extent to which the frequency and layout interaction is indeed exclusively due the right tail of the distribution versus also due a shift in the distribution, for instance by means of Ex-gaussian analyses (e.g. Staub, White, Drieghe, Holloway, & Rayner, 2010). However, such an endeavour would require an experiment with a simpler design (i.e. without a preview manipulation) and a larger number of participants and items to allow for the statistical power necessary for extensive distributional analyses. Nevertheless, we think it is unlikely the interaction between frequency and layout is exclusively due to the right tail across all measurements as our analysis on log-transformed go-past times did show the significant interaction (and see also Veldre et al., 2016[4]).

A somewhat different picture emerges from the single fixation duration data. Besides the effects of frequency and layout (longer fixation times on low frequency and unspaced compared to high frequency and spaced words), an interaction is observed between Preview and Layout such that there is no effect of Preview in the familiar, spaced layout but there is an effect of Preview in unspaced reading. The preview effect is such that the single fixation duration is shorter prior to a non-word preview compared to the normal preview. This parafoveal-on-foveal effect is in the opposite direction as expected, as processing of the non-word preview presumably would increase fixation times. Therefore, we think that a likely explanation for this effect lies in the misspellings being detected from afar, resulting in the point of fixation being drawn towards them (see also Hyönä, 1995). This attraction can result

in result in a short single fixation duration on the pre-boundary word when the saccade towards the post-boundary word falls short and lands on the pre-boundary word, before making a corrective saccade to the intended goal, the post-boundary word. Note that picking up an unusual letter combination this far into the parafovea is compatible not only with a parallel model where all words in the perceptual span gets processed simultaneously up to certain amounts, but also with a serial model such as EZ-Reader (e.g. Reichle, Warren & McConnell, 2009) which incorporates a pre-attentive process in which unusual configurations can be picked up whilst the lexical processing of words is still strictly serial.

Whereas this is only one possible interpretation of this effect, we think the fact that the occurrence of single fixations was actually rarer in the unspaced format (see analysis of number of fixations where this was the only significant effect), might indicate that a single fixation was not the default behaviour in unspaced reading. Instead, perhaps single fixations may sometimes have reflected a mislocated fixation resulting from a saccade that was targeted towards a misspelling. Most importantly, this interpretation can account for this parafoveal-on-foveal effect being limited to single fixation duration, as this would be the measure most influenced by mislocated fixations. Indeed, it would be difficult to give a theoretically plausible interpretation for an effect that is due to the processing of the currently fixated word but that is not present in first fixation duration and neither in gaze but is limited to the single fixation duration on a word.

In previous studies, to examine initial landing sites on words of different lengths, researchers often subdivided the words into five fixation zones (e.g. Rayner, & Fischer, 1996) and this has also been the case with several of the previous unspaced studies (e.g. Rayner et al. 1998; Perea & Acha, 2009). Because we analysed the initial

landing site of the pre-boundary word, which was consistently 5 letters long, we did not have to compensate for different word lengths and analysed the initial landing site as a function of the distance in pixels from the start of the word. Previous studies also mostly reported landing sites as a function of which character the eyes landed on. Again, since we have identical preceding sentences and matched word lengths across our stimuli, we can use the more fine-grained metric of pixels, as this will allow us to pick up more subtle effects in initial landing site compared to when the landing sites are first binned into characters. The saccades landed 6 pixels further to the left into the pre-boundary word when the layout was unspaced compared to when it was spaced. A full character was 11 pixels so the saccade was shortened by about half a character, which is in line with previous findings (e.g. Rayner et al, 1998). As such, we replicate previous observations of the unspaced layout resulting in saccades landing closer to the beginning of the word. However, it is important to note that in the spaced format the space in front of the target word typically gets counted so in terms of which letter of the pre-boundary word the eyes actually land on, the differences between spaced and unspaced layout were almost negligible. This could be due to the current experiment using relatively short words (5 letters long), making it difficult to tease apart differences in initial landing positions. Zang et al. (2013) also observed no differences in initial landing positions when analysing saccade targeting in spaced and unspaced Chinese, potentially also due to the typical short word lengths in Chinese.

The results on the post-boundary word were quite straightforward. For the earliest measures (first and single fixation durations), we observed a main effect of preview and layout such that fixation durations were shorter when the preview of the post-boundary word was identical to the post-boundary word compared to the non-

word preview (the standard preview effect) and of layout (longer fixations times in unspaced compared to spaced format). Unexpectedly, these two main effects were qualified by a three-way interaction between preview, layout and the frequency manipulation on the preceding word. The preview effect was larger when the pre-boundary word was high-frequency and the layout was unspaced compared to the other conditions. This interaction only appeared in the earliest fixation duration measures (first and single fixation duration). We will return to this unexpected interaction in detail in the next paragraph. In skipping rate, gaze duration and go past times only main effects of preview and layout were observed (less skipping and longer fixation times after non-word preview and unspaced format compared to correct preview and spaced format) and in terms of number of fixations the only significant effect observed was more fixations in unspaced compared to spaced layout.

We think the most likely interpretation of the three-way interaction between spacing, preview and the frequency of the pre-target word observed on first and single fixation duration on the post-boundary word lies in the way we created our incorrect previews of the post-target word. The first four letters of the correct preview were substituted such that the resulting non-word preview was orthographically illegal and unpronounceable. In the unspaced format, where determining word boundaries will be quite difficult, it could be that these unusual letter combinations actually afforded a relatively easy way to determine word boundaries. For example, in the sentence "Stephensawasmallsheepkvxmnghiswalkaroundthefield." (the spaced version with the correct preview is "Stephen saw a small sheep during his walk around the field."), the orthographically illegal letter combination "kvxm" may have been visually salient enough to facilitate the insertion of a word boundary after *sheep* and before the illegal string of letters thus allowing for narrowly focussed processing on the pre-boundary

word before a saccade to the post-boundary word. Note how there are no spill-over effects on the post-boundary word in the unspaced layout when there is a non-word preview (if anything the difference after a high- versus a low frequency pre-boundary word goes the opposite direction in first and single fixation duration). We take this as a sign for more narrowly focussed processing of the pre-boundary word when the preview was incorrect, whereas there is a sizeable spill-over effect (30ms in first and 29ms in single fixation duration) when the preview was correct. For the correct preview condition in the unspaced layout, an orthographic aid in the form of an illegal letter combination for determining word boundaries is not present. If we assume that the unspaced layout will invite more distributed processing because the word boundaries need to be determined, a low foveal load (i.e. the high frequency pre-boundary word) results in a larger preview benefit with the fastest fixation durations observed on the post-boundary word in the unspaced layout when the preview was correct and the pre-boundary word was high-frequency. Consequently, a larger preview effect in this condition is observed compared to all the other conditions where some form of word demarcation is offered (spaces or an orthographic illegal string of letters), and a three-way interaction is observed between preview, layout and the frequency of the preceding word. Whereas we acknowledge that this explanation is tentative and needs further investigation (e.g. by running an experiment with orthographically legal, pronounceable non-word previews which would take away the salient orthographic marker), we take the current findings as indicative of a system that in the unspaced layout in all likelihood partakes in more distributed processing to determine word boundaries. However, the system will use any information (e.g. an orthographically illegal string of letters) available to help determine word boundaries which would allow it to more narrowly focus foveal and parafoveal processing.

The use of orthographic cues to determine word boundaries in unspaced layout is compatible with findings in unspaced languages. In Japanese, three character types are used. Kanji, which is a morphographic character set and Hiragana and Katakana, which are syllabaries. Kanji characters are more complex and their visual saliency serves as effective segmentation cues which are used during reading (Sainio, Hyönä, Bingushi, & Bertram, 2007). When reading a mixed Kanji-Hiragana script, Japanese readers show saccade targeting similar to readers in spaced languages (i.e. saccades are targeting positions close to the centre of the word), a behaviour absent from reading in a pure Hiragana script. Readers will also use quite subtle lexical cues to determine word boundaries, for instance the frequency with which certain consonants appear at the beginning or end of a word in Thai to help determine word boundaries and adjust saccade targeting (e.g. Kasisopa, Reilly, Luksaneeyanawin, & Burnham, 2013). Related, in Chinese, Zang et al. (2016) observed that readers used the likelihood of a character being a single character versus the first out of a two character word in word segmentation and modulating the extent of parafoveal processing.

Returning to our predictions, we had anticipated two possible scenarios. In the first scenario, the lack of clear, low-level demarcation would cause lexical processing to be distributed across multiple words to a higher extent than in spaced reading. This would be accompanied by an increase in spill-over and parafoveal-on-foveal effects. In the second scenario, readers would quickly realise there is not much to gain from extensive parafoveal processing, especially since it would result in processing orthographically illegal non-words in certain conditions. As a result, they would adopt a strategy in which they would conservatively try to put a word boundary as early as possible in the stream of characters, after which the system would focus on processing the currently fixated word. Our results for parafoveal-on-foveal effects are mostly in

line with the latter scenario. Even in unspaced reading, we found no parafoveal-on-foveal effects on the pre-boundary word. Only for those cases when there was a single fixation on the unspaced pre-boundary word, an event less common than in the spaced layout, was there a parafoveal-on-foveal effect. However, this parafoveal-on-foveal effect was surprisingly such that shorter single fixations were observed prior to a non-word preview compared to a word preview. As such, we think it is more likely that this effect reflects a few instances where readers picked up the orthographically illegal letter combinations and were targeting saccades to that letter cluster, but fell short and landed at the pre-boundary word. Note that this finding does indicate some attention being allocated far enough in the parafovea for the unusual letter combination to be picked up. However, in the majority of the instances, the patterns in fixation times point towards the processing of the post-boundary preview not influencing the fixation times of the pre-boundary word. As such, we think a fair summary is to describe the observations of parafoveal-on-foveal effects as being quite limited.

For the presence of spill-over effects in the unspaced layout, the nature of the post-target preview was crucial. When the preview consisted of a non-word, the orthographically illegal letter combinations in all likelihood afforded an easy word segmentation and as such a narrow focus on foveal processing, and no spill-over effects were observed. When the post-boundary preview was correct, sizeable spill-over effects were observed. In the latter case, the need to segment the upcoming stream of letters in the words in all likelihood did increase parafoveal processing which led to an increased preview effect when foveal load was low (i.e. high-frequency pre-boundary word) and the preview was correct.

Summarizing, our results point towards a system that tries to deal with the lack of clear word demarcation and the occasional presence of non-words in the parafovea by focusing on processing the currently fixated word. Even though some increased parallel processing (i.e. processing of letters in the parafovea belonging to the next word) seems almost a logical necessity to determine word boundaries, parafoveal-on-foveal effects were mostly absent and spill-over effects were modest in the current experiment. When an orthographic cue afforded by an illegal combination of letters allowed for relatively easy word segmentation, no spill-over effects were observed on the post-boundary word. However, when no such orthographic aid was available increased distributed processing was observed in the unspaced layout. Note that our failure to find an interaction between frequency and preview in the traditional, spaced conditions can be considered a failure to replicate Henderson and Ferreira (1990) who reported a reduced preview benefit after a low-frequency compared to a high-frequency word. Whereas this could be due to our frequency manipulation not being strong enough (we refrained from selecting very low frequency pre-boundary words fearing the words might not be recognised at all in the unspaced format), based on the fact that the sizes of the observed frequency effects on the pre-boundary word in the spaced layout (22 ms in first fixation duration, 20 ms in single fixation duration, 41ms in gaze duration) were completely in line with what can be expected from a successfully implemented frequency manipulation (and numerically larger than the effects of frequency observed in Henderson & Ferreira, 1990), we doubt this very much. This is not the first failure to replicate the Henderson and Ferreira (1990) finding (e.g. see Drieghe, Rayner & Pollatsek, 2005), and we strongly suspect that a meta-analysis carried out on all the studies examining this interaction would result in a significant effect, but one that is relatively small. Occasional null findings can be

expected based on the sampling around a small effect size. Clearly this topic requires further investigation with large sample sizes and meta-analyses of published datasets.

To conclude, in an experiment which examined the effect of spacing on the extent to which lexical processing was distributed across two words, we observed data patterns pointing in the direction of a processing system that is exposed to a situation in which extensive parafoveal processing comes with a cost in an unspaced layout due to occasional non-words in the parafovea. A strategy is adopted by which processing is focused, quite successfully but not perfectly, on the currently fixated word. More distributed processing to determine word boundaries will occur but any orthographic cues that can aid the process of word segmentation (such as the illegal letter combination of the preview) will be used to allow for more focused foveal processing. We believe this hypothesis also to be compatible with the findings reported by Sheridan et al. (2016). They examined parafoveal processing in an experiment in which the spaces were replaced by random numbers and reported a smaller preview effect compared to the spaced layout. The numbers will, at least to a greater extent than in the current experiment where spaces were removed altogether, afford some word length demarcation, and as such, the opportunity to focus processing on the currently fixated word. Therefore, both the numbers in the Sheridan et al. study and the orthographically illegal sequence of letters in the current experiment will provide word segmentation cues that will be used to allow for narrow focused processing of the currently fixated word. Moreover, we also think the use of orthographic cues to assist in determining word boundaries to be compatible with observations from languages that are naturally unspaced such as Chinese, Thai and Japanese (see above).

A great amount of knowledge about foveal versus parafoveal processing and the extent to which readers engage in serial versus parallel lexical processing during

reading has been acquired by means of the influential gaze-contingent boundary paradigm (Rayner, 1975, for a review see Schotter, Angele & Rayner, 2012). A relatively recent development has been to provide novel theoretical insights by using the boundary paradigm to examine serial versus parallel processing in languages that either do or do not afford easy ways to determine word segmentation, and by either artificially adding segmentation to naturally unspaced languages (e.g. Cui et al., 2014) or removing segmentation information from naturally spaced languages (the current study).

REFERENCES

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255-278.

Balota, D.A., Aschenbrenner, A.J., & Yap, M.J. (2013). Additive effects of word frequency and stimulus quality: The influence of trial history and data transformations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 39*, 1563 – 1571.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software, 67*, 1 – 48.

Cui, L., Drieghe, D., Bai, X., Yan, G., & Liversedge, S. P. (2014). Parafoveal preview benefit in unspaced and spaced Chinese reading. *Quarterly Journal of Experimental Psychology, 67*, 2172 – 2188.

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

Drieghe, D., Brysbaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: Does an extra space make a difference? *Vision Research, 45*, 1693-1706.

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.

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

Epelboim, J., Booth, J. R., Ashkenazy, R., Taleghani, A., & Steinman, R. M. (1997). Fillers and spaces in text: The importance of word recognition during reading. *Vision Research, 37,* 2899 - 2914.

Epelboim, J., Booth, J. R., & Steinman, R. M. (1994). Reading unspaced text: Implications for theories of reading eye movements. *Vision Research, 34,* 1735 - 1766.

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.

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.

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

Kasisopa, B., Reilly, R. G., Luksaneeyanawin, S., & Burnham, D. (2013). Eye movements while reading an unspaced writing system: The case of Thai. *Vision Research, 86,* 71 – 80.

Malt, B., & Seamon, J. G. (1978). Peripheral and cognitive components of eye guidance in filled-space reading. *Perception & Psychophysics, 23,* 399-402.

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

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

McGowan, V. A., White, S. J., Jordan, T. R., & Paterson, K. B. (2014). Aging and the use of interword spaces during reading: Evidence from eye movements. *Psychonomic Bulletin & Review, 21*, 740-747.

Perea, M., & Acha, J. (2009). Space information is important for reading. *Vision Research, 49*, 1994-2000.

Perea, M., Tejero, P., & Winskel, H. (2015). Can colours be used to segment words when reading? *Acta Psychologica, 159*, 8-13.

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

R Core Team (2016). R: A language and environment for statistical computing. R foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

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

Rayner, K. (1979). Eye guidance in reading – Fixation locations within words. *Perception, 8,* 21 – 30.

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

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457-1506.

Rayner, K. (2014). The gaze-contingent moving window in reading Development and review. *Visual Cognition, 22*, 242 – 258.

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

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movments during reading and scanning are different. *Perception & Psychophysics, 58*, 734 – 747.

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, & Pollatsek, A. (1996). Reading unspaced text is not easy: Comments on the implications of Epelboim et al.s (1994) study for models of eye movement control during reading. *Vision Research, 36*, 461 – 470.

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., Warren, T., & McConnell, K. (2009). Using EZ Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review, 16*, 1-21.

Sainio, M., Hyönä, J., Bingushi, K. & Bertram, R. (2007). The role of interword spacing in reading Japanese: An eye movement study. *Vision Research,* 47, 2575 – 2584.

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

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception & Psychophysics, 74*, 5 - 35.

Sheridan, H., Rayner, K., & Reingold, E. M. (2013). Unsegmented text delays word identification: Evidence from a survival analysis of fixation durations. *Visual Cognition, 21*, 38-60.

Sheridan, H., Reichle, E. D., & Reingold, E. M. (2016). Why does removing inter-word spaces produce reading deficits? The role of parafoveal processing. *Psychonomic Bulletin & Review*, *23,* 1543 – 1552.

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

Slattery, T. J., & Rayner, K. (2013). Effects of intraword and interword spacing on eye movements during reading: Exploring the optimal use of space in a line of text. *Attention, Perception, & Psychophysics, 75*, 1275-1292.

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

Starr, M. S., & Rayner, K. (2001). Eye movements during reading : some current controversies. *Trends in Cognitive Sciences, 5*, 156 – 163.

Staub, A., White, S. J., Drieghe, D., Hollway, E. C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of*

Experimental Psychology: Human Perception and Performance, 36, 1280–1293.

Van Heuven, W.J.B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67,* 1176 – 1190.

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N + 1 and word N + 2 during reading: A critical review and Bayesian meta-analysis. *Psychonomic Bulletin & Review*, Advance online publication.

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

Veldre, A., Drieghe, D., & Andrews, A. (2017). Spelling ability selectively predicts the magnitude of disruption in unspaced text reading. *Journal of Experimental Psychology: Human Perception and Performance. In Press.*

Yan, M., Richter, E.D., Shu. H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin & Review, 16,* 561-566.

Yang, S.-N., & McConkie, G.W. (2001). Eye movements during reading: a theorie of saccade initiation times. *Vision Research, 41,* 3567 – 3585.

Yen, M., Radach, R., Tzeng, O., Hung, D., & Tsai, J. (2009). Early parafoveal processing in reading Chinese sentences. *Acta Psychologica, 131,* 24–33.

Yu, L., Cutter, M., Yan, G., Xuejun, B., Fu, Y., Drieghe, D., & Liversedge, S.P. (2016). Word n+2 preview effects in three-character Chinese idioms and phrases. *Language, Cognition, and Neuroscience, 31,* 1130 – 1149.

Zang, C., Liang, F., Bai, X., Yan, G., & Liversedge, S.P. (2013). Interword spacing and landing position effects during Chinese reading in children and adults. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 720 – 734.

Zang, C., Liversedge, S.P., Bai, X., & Yan, G. (2011). Eye movements during Chinese reading. In Liversedge, S.P., Gilchrist, I., & Everling, S. (Eds.), *The Oxford Handbook of Eye Movements* (pp. 961-978). Oxford University Press.

Zang, C., Wang, Y., Bai, X., Yan, G., Drieghe, D., & Liversedge, S.P. (2016). The use of probabilistic lexicality cues for word segmentation in Chinese reading. *The Quarterly Journal of Experimental Psychology, 69*, 548 – 560.

Footnotes.

1. Display change experiments routinely remove about 10-15% of trials due to the exclusion criteria that were also implemented for the current analysis. Our higher exclusion rate is in all likelihood due to a combination of very strict inclusion criteria and the unspaced nature of half of our trials. By removing the space in front of the post-boundary word and by saccades landing earlier in the target word in the unspaced compared to the spaced conditions (see analyses below), there will be comparatively more instances where the eyes land closer to the boundary, resulting in less time for the tracker to detect samples across the boundary and execute the display change fast enough, and as such more instances of the events which are typically excluded (e.g. late display changes). This is consistent with some reports of comparatively high exclusion rates in display change experiments in unspaced languages such as Chinese (e.g. Cui et al. 2014 removed participants to reduce the exclusion rate, Yu, et al., 2016 reported exclusion rates of 27-35%).

2. In eye-tracking during reading experiments which employ the gaze-contingent boundary paradigm, a consensus has gradually developed in recent years to ask participants after the experiments whether they noticed any of the display changes. Participants who notice too many display changes (e.g. more than 5 display changes in a reasonably large scale study, Veldre & Andrews, 2016) are removed due to concerns that these participants might change their reading strategy when noticing these changes, resulting in behaviour that cannot be considered normal reading. In the current experiment, the majority of participants noticed display changes, likely due to nature of the manipulations. Whereas we fully agree that this could have led to a change in reading strategy,

we would like to point out that this criticism applies equally to the entire unspaced reading paradigm. The goal of the current study was to see how eye movement behaviour and lexical processing adjusts to an unnatural reading situation.

3. For the eye movement measures on the pre-boundary word, the random structure for the LMM of the first fixation durations was (1 + layout + preview + frequency | Subject) + (1 + layout + preview + frequency | Stimuli), for single fixation duration it was (1 + preview + frequency| Subject) + (1 + layout + preview| Stimuli), for gaze duration it was (1 + layout + preview + frequency | Subject) + (1 + layout + preview + frequency | Stimuli), for go-past times it was (1 + layout + preview + frequency | Subject) + (1 + layout + preview + frequency | Stimuli), and for skipping rate, number of fixations and landing sites it was (1| Subject) + (1| Stimuli). For the eye movement meansures on the post-boundary word, the random structure for the LMM for first fixation duration was (1 + layout + preview + frequency | Subject) + (1 | Stimuli), for single fixation duration was (1 + layout + preview + frequency | Subject) + (1 + layout + preview + frequency | Stimuli), for gaze duration was (1 + layout + preview + frequency | Subject) + (1 + layout + preview + frequency | Stimuli), for go-past times was (1 + preview + frequency + layout | Subject) + (1 | Stimuli), for skipping rate was (1| Subject) and for the number of fixations it was (1 | Subject) + (1 | Stimuli).

4. Veldre, Drieghe and Andrews (2017) analysed their spacing and frequency experiment with LMM's carried out on log-transformed data. They did report an interaction indicating a higher frequency effect in unspaced reading compared to standard spacing in gaze duration. However, this was in the

context of an individual differences study where factors were included in the model that were not available in the current study (e.g., reading ability and spelling ability of the participants).

Table 1. An example sentence from the experiment illustrating each of the 8 conditions.

1. High frequency noun – correct preview - spaced

On stage there was a young child *singing* with the most amazing voice.

2. High frequency noun – incorrect preview - spaced

On stage there was a young child *tuvqing* with the most amazing voice.

3. Low frequency noun – correct preview - spaced

On stage there was a young tenor *singing* with the most amazing voice.

4. Low frequency noun – incorrect preview - spaced

On stage there was a young tenor *tuvqing* with the most amazing voice.

5. High frequency noun – correct preview - unspaced

Onstagetherewasayoungchild*singing*withthemostamazingvoice.

6. High frequency noun – incorrect preview - unspaced

Onstagetherewasayoungchild*tuvqing*withthemostamazingvoice.

7. Low frequency noun – correct preview - unspaced

Onstagetherewasayoungtenor*singing*withthemostamazingvoice.

8. Low frequency noun – incorrect preview - unspaced

Onstagetherewasayoungtenor*tuvqing*withthemostamazingvoice.

Note: The stimuli shown in italics indicate the preview for each condition prior to the eyes' crossing of the display change boundary. The preview was always replaced by the correct word after the boundary had been crossed.

Table 2. Mean (and Standard Deviation) Eye Movement Measures for the Pre-Boundary Word as a function of Layout, Preview Condition of the Post-Boundary Word and Frequency of the Pre-Boundary Word.

	Spaced				Unspaced			
	Correct Preview		Non-word Preview		Correct Preview		Non-word Preview	
	High Frequency	Low Frequency	High Frequency	Low Frequency	High Frequency	Low Frequency	High Frequency	Low Frequency
First Fixation Duration (ms)	213 (62)	230 (62)	222 (70)	239 (78)	307 (102)	326 (119)	295 (114)	319 (114)
Single Fixation Duration (ms)	208 (59)	229 (64)	221 (68)	239 (80)	296 (94)	333 (113)	278 (106)	299 (106)
Gaze Duration (ms)	220 (73)	258 (96)	226 (74)	270 (115)	365 (171)	470 (247)	370 (207)	447 (272)
Go Past Times (ms)	239 (105)	290 (141)	250 (106)	282 (128)	425 (243)	652(490)	519 (357)	716 (620)
Skipping Rate (%)	23 (42)	16 (37)	21 (41)	18 (39)	18 (39)	8 (27)	13 (33)	11 (31)
Number of Fixations	1.00 (.00)	1.16 (.37)	1.00 (.00)	1.12 (.33)	1.21 (.41)	1.39 (.55)	1.20 (.40)	1.36 (.61)
Initial Landing Site (pixels)	27 (16)	27 (18)	31 (17)	28 (15)	22 (13)	21 (13)	23 (13)	22 (13)

Table 3. Results of the (Generalized) Linear Mixed-Effects Models for the Pre-Boundary Word Analyses. Significant Effects are indicated in Bold.

Measure	Fixed effect	b	SE	t/z
First Fixation Duration	Intercept	5.53	0.03	**184.28**
	Preview	-0.00	0.02	-0.11
	Layout	0.31	0.03	**11.99**
	Frequency	0.07	0.02	**3.18**
	Preview x Layout	-0.06	0.04	-1.46
	Preview x Frequency	0.03	0.04	0.70
	Layout x Frequency	-0.03	0.04	-0.67
	Preview x Layout x Frequency	0.03	0.09	0.37
Single Fixation Duration	Intercept	5.52	0.03	**194**
	Preview	-0.03	0.03	-1.04
	Layout	0.29	0.03	**10.70**
	Frequency	0.09	0.03	**3.49**
	Preview x Layout	-0.14	0.05	**-2.88**
	Preview x Frequency	-0.02	0.05	-0.32
	Layout x Frequency	-0.00	0.05	-0.01
	Preview x Layout x Frequency	-0.02	0.10	-0.17
Gaze Duration	Intercept	5.67	0.04	**133.61**
	Preview	-0.02	0.03	-0.64
	Layout	0.47	0.03	**14.33**
	Frequency	0.17	0.03	**5.08**
	Preview x Layout	-0.07	0.05	-1.36

	Preview x Frequency	-0.02	0.05	-0.31
	Layout x Frequency	0.02	0.05	0.43
	Preview x Layout x Frequency	-0.05	0.11	-0.42
	Intercept	5.82	0.05	**114.03**
	Preview	0.05	0.04	1.17
Go Past Times	Layout	0.64	0.05	**12.03**
	Frequency	0.22	0.04	**4.85**
	Preview x Layout	0.06	0.07	0.96
	Preview x Frequency	-0.6	0.07	-0.89
	Layout x Frequency	0.14	0.07	**2.05**
	Preview x Layout x Frequency	0.01	0.13	0.07
	Intercept	-2.05	0.23	**-9.11**
	Preview	0.01	0.19	0.04
Skipping Rate	Layout	-0.67	0.19	**-3.49**
	Frequency	-0.51	0.19	**-2.66**
	Preview x Layout	0.01	0.38	0.03
	Preview x Frequency	0.60	0.38	1.57
	Layout x Frequency	-0.22	0.38	-0.59

	Preview x Layout x Frequency	0.44	0.76	0.57
Number of Fixations	Intercept	0.16	0.03	**4.95**
	Preview	-0.02	0.06	-0.27
	Layout	0.19	0.06	**2.90**
	Frequency	0.13	0.06	**2.06**
	Preview x Layout	0.00	0.13	0.01
	Preview x Frequency	-0.02	0.13	-0.18
	Layout x Frequency	0.00	0.13	0.02
	Preview x Layout x Frequency	0.03	0.26	0.10
Initial Landing Site	Intercept	25.16	0.54	**46.64**
	Preview	1.43	1.02	1.40
	Layout	-5.96	1.02	**-5.84**
	Frequency	-1.16	1.02	-1.14
	Preview x Layout	-0.81	2.04	-0.40
	Preview x Frequency	-2.11	2.04	-1.03
	Layout x Frequency	0.25	2.04	0.12
	Preview x Layout x Frequency	2.20	4.08	0.54

Table 4. Mean (and Standard Deviation) Eye Movement Measures for the Post-Boundary Word as a function of Layout, Preview Condition of the Post-Boundary Word and Frequency of the Pre-Boundary Word.

	Spaced				Unspaced			
	Correct Preview		Non-word Preview		Correct Preview		Non-word Preview	
	High Frequency	Low Frequency	High Frequency	Low Frequency	High Frequency	Low Frequency	High Frequency	Low Frequency
First Fixation Duration (ms)	222 (51)	224 (66)	253 (72)	262 (92)	262 (95)	292 (102)	330 (109)	314 (97)
Single Fixation Duration (ms)	218 (49)	219 (55)	255 (64)	264 (78)	253 (84)	282 (104)	324 (96)	320 (107)
Gaze Duration (ms)	257 (109)	258 (107)	297 (116)	306 (114)	363 (202)	405 (213)	436 (197)	431 (207)
Go Past Times (ms)	275 (127)	282 (127)	344 (159)	339 (133)	538 (441)	714 (596)	715 (516)	862 (767)
Skipping Rate (%)	18 (39)	19 (39)	9 (29)	6 (24)	7 (25)	5 (22)	2 (13)	5 (22)
Number of Fixations	1.12 (.32)	1.17 (.38)	1.18 (.39)	1.23 (.42)	1.35 (.52)	1.42 (.63)	1.40 (.61)	1.41 (.62)

Table 5. Results of the (Generalized) Linear Mixed-Effects Models for the Post-Boundary Word Analyses. Significant Effects are indicated in Bold.

Measure	Fixed effect	b	SE	t/z
First Fixation Duration	Intercept	5.54	0.03	**219.53**
	Preview	0.15	0.02	**6.72**
	Layout	0.21	0.03	**8.17**
	Frequency	0.02	0.02	0.87
	Preview x Layout	0.02	0.04	0.55
	Preview x Frequency	-0.06	0.04	-1.39
	Layout x Frequency	0.02	0.04	0.52
	Preview x Layout x Frequency	-0.19	0.08	**-2.27**
Single Fixation Duration	Intercept	5.53	0.02	**244.94**
	Preview	0.18	0.02	**7.15**
	Layout	0.20	0.03	**5.95**
	Frequency	0.02	0.03	0.79
	Preview x Layout	0.02	0.05	0.51
	Preview x Frequency	-0.04	0.05	-0.98
	Layout x Frequency	0.02	0.05	0.51
	Preview x Layout x Frequency	-0.20	0.09	**-2.17**
Gaze Duration	Intercept	5.73	0.04	**140.94**
	Preview	0.17	0.03	**5.53**
	Layout	0.35	0.04	**9.11**
	Frequency	0.02	0.04	0.56

	Preview x Layout	-0.03	0.05	-0.54
	Preview x Frequency	-0.04	0.05	-0.72
	Layout x Frequency	0.00	0.05	0.06
	Preview x Layout x Frequency	-0.15	0.10	-1.48
Go Past Times	Intercept	5.97	0.05	**118.20**
	Preview	0.23	0.04	**5.71**
	Layout	0.66	0.06	**11.37**
	Frequency	0.08	0.04	**2.00**
	Preview x Layout	0.02	0.07	0.30
	Preview x Frequency	-0.08	0.07	-1.17
	Layout x Frequency	0.13	0.07	1.86
	Preview x Layout x Frequency	-0.11	0.14	-0.78
Skipping Rate	Intercept	-3.06	0.30	**-10.32**
	Preview	-0.90	0.29	**-3.15**
	Layout	-1.25	0.29	**-4.33**
	Frequency	0.08	0.29	0.28

	Preview x Layout	0.43	0.57	0.74
	Preview x Frequency	0.51	0.57	0.90
	Layout x Frequency	0.68	0.57	1.19
	Preview x Layout x Frequency	1.94	1.14	1.69
Number of Fixations	Intercept	0.25	0.03	**8.37**
	Preview	0.03	0.06	0.58
	Layout	0.17	0.06	**2.95**
	Frequency	0.04	0.06	0.60
	Preview x Layout	-0.03	0.12	-0.27
	Preview x Frequency	-0.03	0.12	-0.22
	Layout x Frequency	-0.01	0.12	-0.11
	Preview x Layout x Frequency	-0.03	0.24	0.15

Figure Captions

Figure 1. Effect display for the significant interaction of Preview and Layout for Single Fixation Durations on the Pre-Boundary Word. A 95-percent confidence interval is drawn around the estimated effect.

Figure 2. Effect display for the significant interaction of Preview and Frequency for Go-Past Times on the Pre-Boundary Word. A 95-percent confidence interval is drawn around the estimated effect.

Figure 3. Effect display for the significant three-way interaction of Preview, Spacing and Layout for First Fixation Durations on the Post-Boundary Word. A 95-percent confidence interval is drawn around the estimated effect.

Figure 4. Effect display for the significant three-way interaction of Preview, Spacing and Layout for Single Fixation Durations on the Pre-Boundary Word. A 95-percent confidence interval is drawn around the estimated effect.
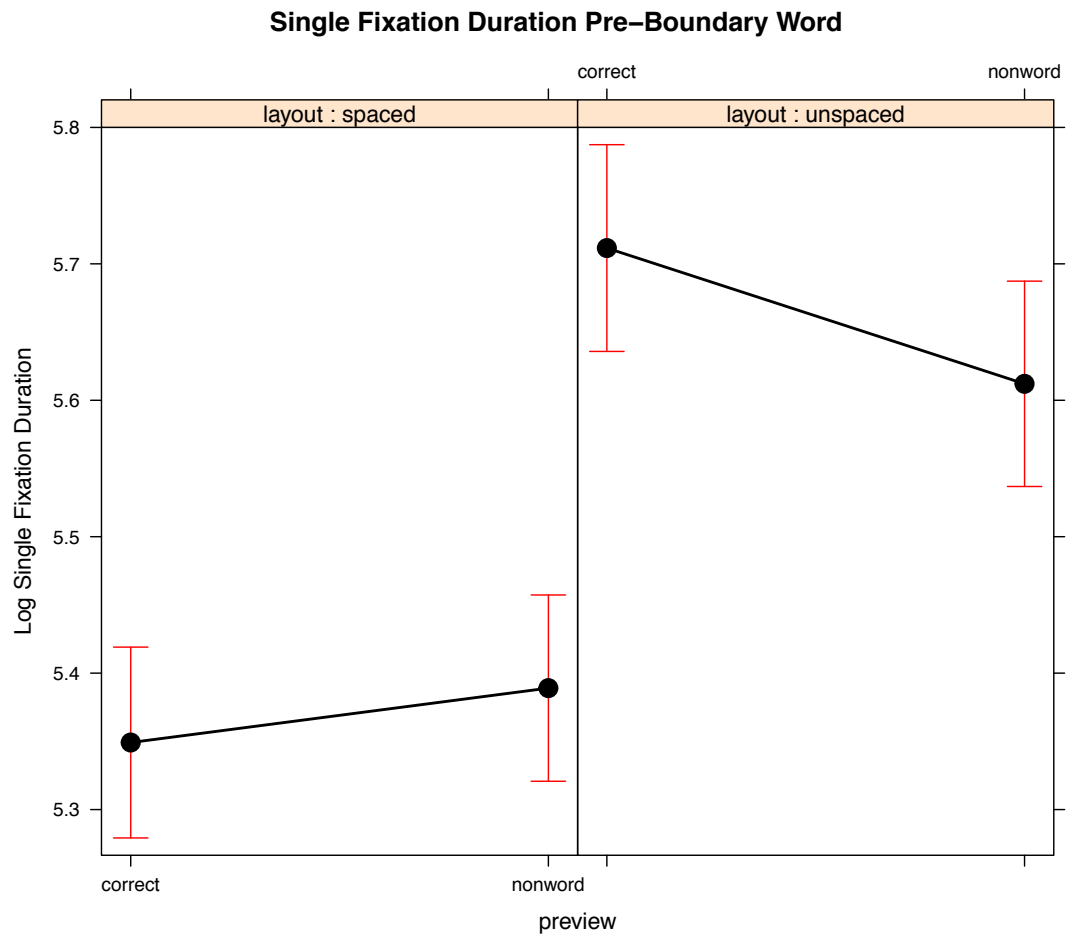
Figure 1

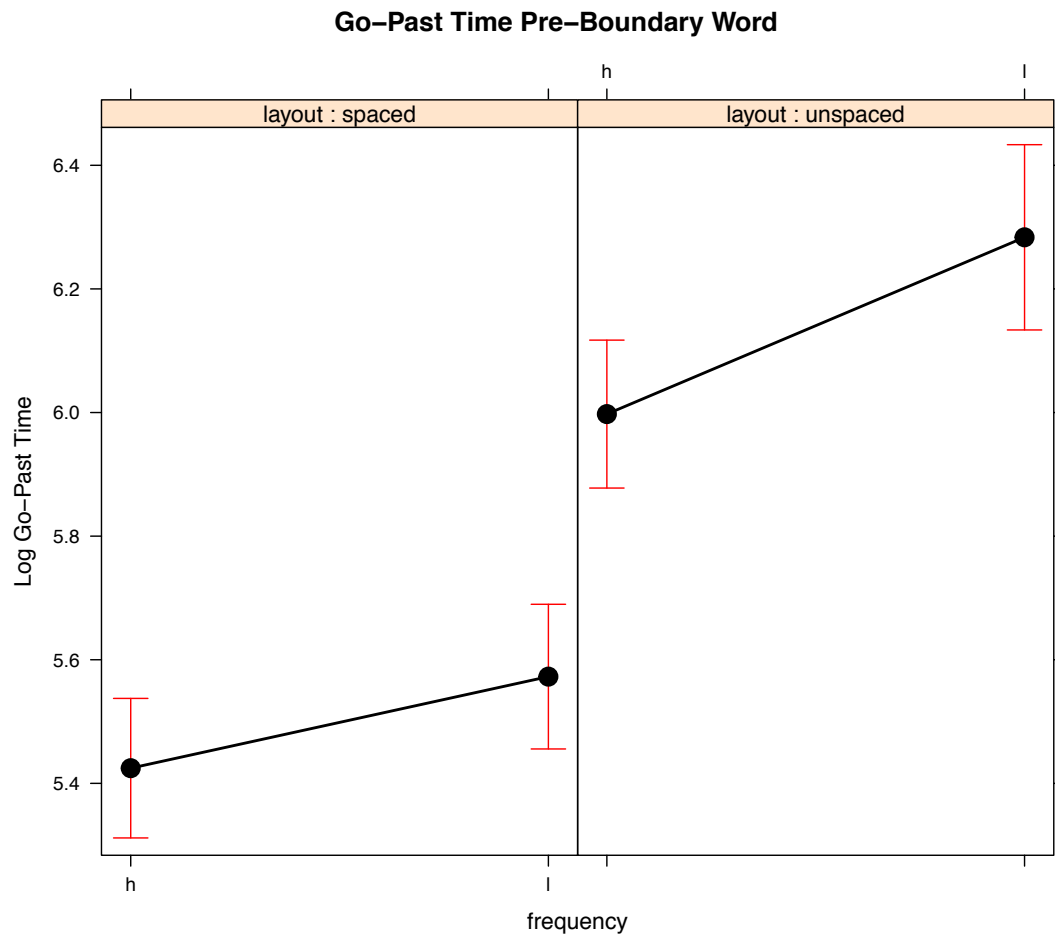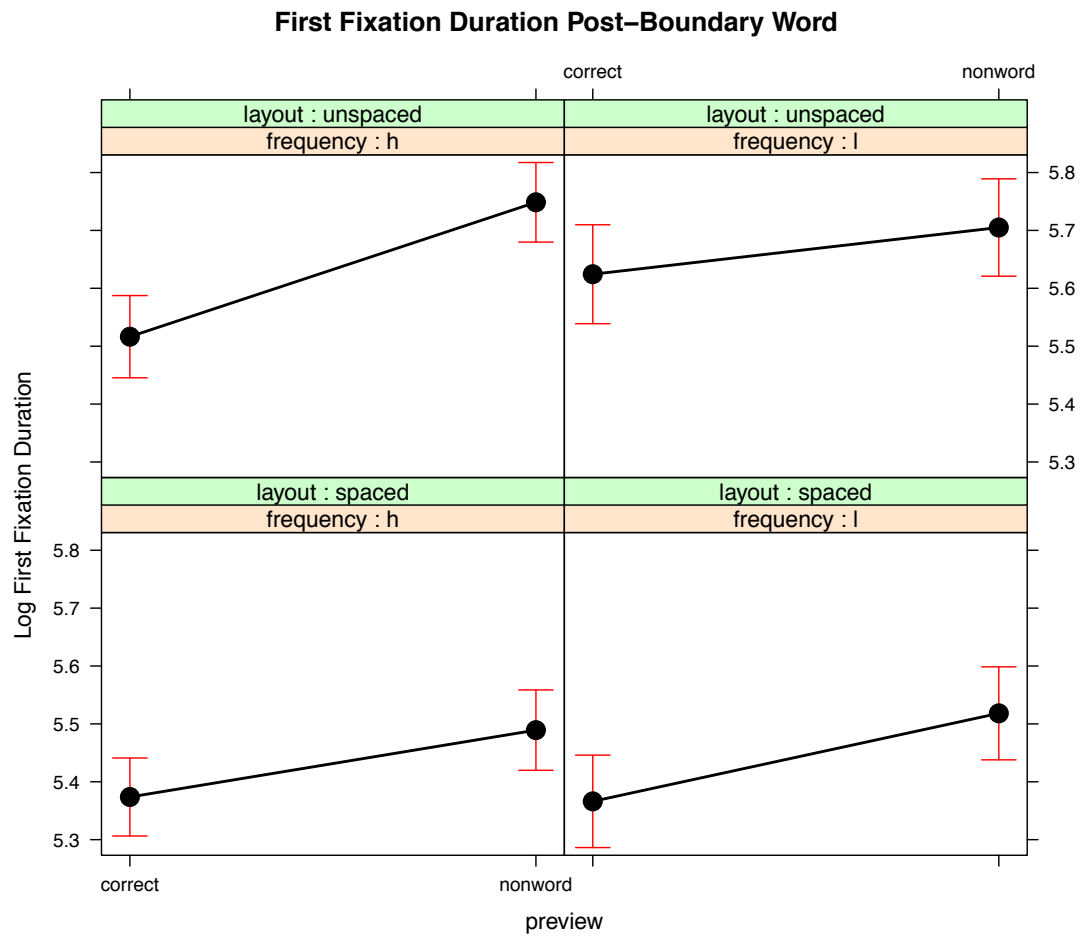**Single Fixation Duration Pre−Boundary Word**

Figure 2

Go−Past Time Pre−Boundary Word

Figure 3

First Fixation Duration Post–Boundary Word

Figure 4

Single Fixation Duration Post−Boundary Word