Manuscript currently in press at the Journal of Experimental Psychology: Human Perception and Performance

Parafoveal-foveal Overlap Can Facilitate Ongoing Word Identification During Reading: Evidence from Eye Movements

Bernhard Angele, Randy Tran, and Keith Rayner

Department of Psychology

University of California, San Diego

Correspondence to:

Bernhard Angele

Department of Psychology

University of California, San Diego

La Jolla, CA 92093

Email: bangele@ucsd.edu

Running Head: Parafoveal-foveal overlap in reading

Abstract

Readers continuously receive parafoveal information about the upcoming word in addition to the foveal information about the currently fixated word. Previous research (Inhoff, Radach, Starr, & Greenberg, 2000) showed that the presence of a parafoveal word which was similar to the foveal word facilitated processing of the foveal word. In three experiments, we used the gaze-contingent boundary paradigm (Rayner, 1975) to manipulate the parafoveal information that subjects received before or while fixating a target word (e.g. *news*) within a sentence. Specifically a reader's parafovea could contain a repetition of the target (*news*), a correct preview of the post-target word (*once*), an unrelated word (*warm*), random letters (*cxmr*), a nonword neighbor of the target (*niws*), a semantically related word (*tale*), or a nonword neighbor of that word (*tule*). Target fixation times were significantly lower in the parafoveal repetition condition than in all other conditions, suggesting that foveal processing can be facilitated by parafoveal repetition. We present a simple model framework that can account for these effects.

It is obvious that listeners receive linguistic information one word at a time and that, in terms of lexical processing, speech perception is a serial process (Rayner & Clifton, 2009). In contrast, during reading, multiple words are visually available to the reader simultaneously. But, do readers use information from multiple words in parallel? There is considerable evidence that readers can pre-process parafoveal words before fixating them and this leads to faster processing of the word once it is fixated (see Rayner, 1998, 2009; Schotter, Angele, & Rayner, 2012). The issue of serial versus parallel processing is also at the heart of debates and models concerning eye movement control in reading. According to the E-Z Reader model (Pollatsek, Reichle, & Rayner, 2006; Rayner, Li, & Pollatsek, 2007; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006; Reichle, Warren, & McConnell, 2009), lexical processing is serial. Although there are some parallel components in the model, such as the overlapping saccade planning and word identification stages or the parallel identification of letters within a word, in E-Z Reader the upcoming word $n+1$ is not lexically processed until the processing of the currently fixated word n has achieved lexical access. In contrast, in models like SWIFT (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005) and Glenmore (Reilly & Radach, 2003, 2006) multiple words can in principle be processed in parallel though the fixated word generally receives the most activation in these models. The question of the plausibility of parallel lexical processing of words during reading is highly controversial (see Reichle, Liversedge, Pollatsek, & Rayner, 2009) and beyond the scope of the present article. Rather, we address a somewhat different issue, namely the extent to which parafoveal word information influences processing of the foveal fixated word.

This latter issue has typically been couched in terms of parafoveal-on-foveal effects and here too the findings are rather controversial (see Schotter et al., 2011). Parafoveal-on-foveal effects refer to the possibility that the characteristics of word *n+1* influence the duration of the fixation on word *n*. There are data suggesting that unusual orthographic features of word *n+1* slow processing of word *n* (Blanchard, Pollatsek, & Rayner, 1989; White & Liversedge, 2006, 2004). Such effects are also known as *orthographic parafoveal-on-foveal effects*. The data concerning lexical processing are as already noted controversial, with some studies showing no effects of the frequency of word *n+1* on word *n* (Henderson & Ferreira, 1990; Rayner, Fischer, & Pollatsek, 1998) and others showing effects (Kliegl, Nuthmann, & Engbert, 2006). However, such effects are typically observed in corpus-based analyses (where all words in the sentence are analyzed, rather than controlled target-word experiments which typically do not show an effect). Furthermore, the direction of the effect is not consistent experimental studies, with some yielding shorter fixations on word *n* when word *n+1* is frequent (e.g. Kennedy, 1998) and some showing longer fixations on word *n* when word *n+1* is frequent (e.g. Kennedy, 2000; Kennedy, Pynte, & Ducrot, 2002). This inconsistency was investigated further by Hyönä and Bertram (2004), who argued that it might be explained by parafoveal "magnetic attraction", that is, the tendency of subjects to move their eyes towards unusual information in the parafovea.

In the present paper, we address a slightly different issue: can orthographic information from word *n+1* facilitate the processing of word *n* as well as disrupt it? As noted above, unusual parafoveal orthographic information typically slows reading. However, a study by Inhoff, Radach, Starr, and Greenberg (2000) suggests that orthographic information might be facilitatory. They had subjects read sentences in which, in the critical condition, there was an orthographic repetition. Thus, subjects read sentences such as:

1a. Did you see the picture of her mother's mother at the meeting. (repetition)

1b. Did you see the picture of her mother's father at the meeting. (related)

1c. Did you see the picture of her mother's garden at the meeting. (unrelated)

Inhoff et al. observed shorter fixation times on *mother's* in the repetition and related conditions than in the unrelated condition. While this finding is quite interesting, it is also the case that the syntactic structure was somewhat unusual and, furthermore, the repetition condition may have been salient to readers.

In two experiments, we used the gaze contingent boundary paradigm (Rayner, 1975) to present parafoveally repeated words in more natural reading materials. In the boundary paradigm, readers' eye movements are monitored and, when a readers' gaze crosses an invisible boundary location, a preview (that is, pre-display change) word (usually the word to the right of the boundary) is replaced by a post-display change word. Because the display change occurs during a saccade, when vision is suppressed (Matin, 1974), readers are not typically aware of the change. This enabled us to insert parafoveal repetition previews into our stimuli without readers becoming aware of their presence. As in other boundary experiments, we then examined the amount of time during which readers fixated pre-boundary (target) and the post-boundary (post-target) word as a function of what the preview for the post-boundary word had been. In Experiment 1, there were four conditions as shown in Figure 1.

In the identical control condition, although there was a display change, the preview was changed to an identical post-change word. In the repetition condition, while readers were looking at the target word *news*, the preview was also *news*, while in the unrelated condition it was *warm* and in the random letter condition it was a string of unrelated letters. As soon as the readers' eyes crossed the boundary location, the preview changed to the post-target word *once*.

In Experiment 2, we investigated whether any facilitation associated with repetition was mediated by orthographic or semantic information by adding previews that shared either orthographic (*niws*) or semantic (*tale*) properties with the target word.

Experiment 1

Method

Subjects.

Forty University of California, San Diego students participated in this experiment for course credit. All were native speakers of English, had either normal or corrected to normal vision, and were naïve concerning the purpose of the experiment.

Apparatus.

An SR Research Eyelink 1000 eyetracker was used to record subjects' eye movements with a sampling rate of 1000 Hz. Subjects read sentences displayed on an Iiyama Vision Master Pro 454 video monitor with a refresh rate of 150 Hz. Viewing was binocular, but only the right eye was recorded. Viewing distance was approximately 60 cm, with 3.8 letters equaling one degree of visual angle. A monospace font (Courier New) was used to ensure that all words of the same length had the same width on the screen.

Materials.

INSERT FIGURE 1 ABOUT HERE

One hundred and twenty experimental sentences were generated for the experiment. Each sentence contained a target word which was followed by a post-target word of the same length. Using the gaze-contingent boundary paradigm (Rayner, 1975), we manipulated the information available from the parafovea (i.e. the preview of the post-target word) while subjects were fixating the target word (see Figure 1). For example, while subjects were fixating the target word *news* in the sentence *Victor read the news once this morning,* the parafoveal preview of the post-target word consisted of (1) the actual post-target word (*once*, identical control condition), (2) a repetition of the target word (*news*), (3) an unrelated word (*warm*), or (4) a random letter string (*rzmc*). Once subjects moved their eyes across the boundary located to the right of the target word, the post-target word preview changed to the actual post-target word (*once*). Table 1 shows word frequency (from the Corpus of Contemporary American English, Davies, 2011), and mean log bigram frequency (from the *N-Watch* software, Davis, 2005) for the preview.

INSERT TABLE 1 ABOUT HERE

Procedure.

Once the readers moved their eyes across the boundary, the preview was replaced by the actual post-target word (e.g. *once*). Custom-designed software (*Eyetrack*, developed at the University of Massachusetts at Amherst) was used to maximize the chance that the display changes were completed before the end of the saccades that triggered them. In particular, the *Eyetrack* software is able to rapidly update the screen by writing directly to the buffer of the graphics card. Display change delays are estimated taking the screen refresh rate (150 Hz at a resolution of 1024 by 768) into account. Approximately 33% of the sentences were followed by a two-alternative forced choice comprehension question which subjects answered by pressing the button corresponding to the correct answer on a button box. The mean accuracy was greater than 85% for all subjects.

Results

From the eye tracking data, we computed the following fixation time measures on three target regions, namely the pre-target word, the target word, and the post-target word: first fixation duration (FFD), single fixation duration (SFD), gaze duration (GD), go-past time (go-past), total viewing time (TVT), landing position, fixation probability, and the probability of making a regression out of the word. First fixation duration is the mean duration of the first fixation on a word, regardless of whether there are subsequent fixations on that word or not, and is considered a measure of early processing (Rayner, 1998, 2009). Single fixation duration is computed in the same way as first fixation duration, but it only takes into account cases in which a word was fixated exactly once during first pass. Mean gaze duration is the sum of the duration of the first fixation on a word and the durations of all subsequent refixations before leaving the word. While it is still a measure of early processing, it can capture some later processing difficulties that force a reader to refixate on a word. Mean go-past time includes all the fixations used to calculate gaze duration, but additionally considers the durations of fixations that are made to the left of the word in question from the time a reader first enters that word from the left until the reader leaves the word to the right. As such, it is sensitive to integration difficulties that require regressions. Finally, total viewing time is the sum of the durations of all fixations on a word during a trial, regardless of whether they occurred during first pass or later. In addition to the fixation time measures, we computed the probability of subjects making a first-pass fixation on and a regression out of each target region. It has been shown that easier to process words are more likely to be skipped (e.g. Brysbaert, Drieghe, & Vitu, 2005; for reviews see Rayner, 1998, 2009 and Schotter et al., 2012). On the other hand, difficulties in integrating the current or a previous word into the sentence context can cause readers to make a regression out of a word. Go-past time and regression probability are closely related. Finally, the initial landing position on a word is mostly determined by word length, but landing positions towards the beginning of a word can be interpreted as being indicative of processing difficulties and a cautious saccadic strategy as well (Rayner et al., 1998). Tables 2a through 4a show means and standard deviations for each condition and dependent measure.

We took great care to minimize any possible artifacts caused by the display change technique. After the experiment, subjects were asked whether they had noticed anything unusual during the experiment. If they did not mention the display changes, they were asked specifically whether they had observed any display changes. If subjects indicated that they had seen more than five display changes, their data were excluded from the experiment. This criterion led to the exclusion of 7 subjects out of 47 subjects which originally participated in the experiment for a total of 40 subjects included in the analysis. Since late or improper display changes affect fixation times on the target words (Slattery, Angele, & Rayner, 2011; White, Rayner, & Liversedge, 2005), we eliminated trials with track losses or display changes that completed more than 9 ms after fixation onset as well as trials in which a blink occurred immediately before or during a fixation on the target word (23 % of the data)[1] from the post-target word analyses. As the display changes only occurred after all first-pass fixations on the pre-target and target words had ended, display change issues did not affect fixation times on those words. Because of this, we applied a much more lenient criterion to the analyses of the pre-target and target words, only eliminating trials if a display change was more than 25 ms late (which would indicate massive problems on a trial). As a consequence, only 2.73 % of trials were eliminated from the pre-target and target word analyses. Since total viewing times contain second-pass fixations and might be affected by display change problems, we used the stricter elimination criterion for all target words for that measure. Finally, if a fixation was shorter than 80 ms and located within one character space (11 pixels) of another fixation, it was merged into that fixation. All other fixations shorter than 80 ms or fixations longer than 800 ms (less than 1% of the data) were deleted.

We report inferential statistics based on linear mixed models (LMM) with subjects and items as crossed random effects (Baayen, Davidson, & Bates, 2008). In order to fit the LMMs, the lmer function from the lme4 package (Bates, Maechler, & Dai, 2009) was used within the R Environment for Statistical Computing (R Development Core Team, 2011). For each factor, we report regression coefficients (b), standard errors, and t-values. We do not report p-values, since it is not clear how to determine the degrees of freedom for LMMs, making it difficult to estimate p-values. However, since our analyses contain a large number of subjects and items and only a few fixed and random effects are estimated, we can assume that the distribution of the t-values estimated by the LMMs approximates the normal distribution. We will therefore use the two-tailed criterion $|t| \geq 1.96$ which corresponds to a significance test at the 5% α-level.

The analysis included preview (identical control vs. repetition vs. unrelated word vs. nonword) as a fixed effect as well as random intercepts for subjects and items. As we were mostly interested in the differences between each of the experimental preview conditions and the control condition, we used treatment contrasts with the identical control condition (e.g. *once*) as the baseline. As a result, the t-value for each of the contrasts indicates if a dependent measure in one of the experimental preview conditions is significantly different from the control condition, while the b-values reflect the magnitude of the difference in ms. For fixation probability and probability of regressions out, logistic LMMs were used (Gelman & Hill, 2007). The b-values are therefore not as easily interpretable as for the fixation time analyses. Tables 2b, 3b, and 4b show the results for all models fitted on the pre-target, target, and post-target. We will now discuss the effects on each word in detail.

INSERT TABLE 2 ABOUT HERE

Pre-target word.

As shown in Figure 1, the pre-target word was always located at least five characters (the length of the target word plus the preceding space) to the left of the boundary and the preview. Because of this, we did not expect to find strong effects of the preview manipulation on any of the dependent variables on the pre-target word. Both the means pattern (Table 2a) and the LMMs (Table 2b) confirmed these expectations: None of the predictors reached significance, with one exception: Total viewing times in the unrelated and in the nonword condition were significantly longer than in the control condition. Since there were no differences in first-pass time, this suggests that the unrelated and nonword previews lead subjects to return to the pre-target word and re-read it more often than the control condition. None of the other measures showed significant effects of the preview manipulation.

INSERT TABLE 3 ABOUT HERE

Target word.

Given Inhoff et al.'s (2000) results, we expected to observe shorter first-pass fixation times in the parafoveal repetition condition compared to the identical control condition. The means (Table 3a) and the LMM results (Table 3b) show that in all fixation time measures except total viewing time, fixations in the repetition condition were indeed significantly shorter than in the control condition, indicating that we succeeded in replicating Inhoff et al.'s (2000) results. In first fixation duration, and go-past time, there was no difference between either the unrelated or the nonword condition and the control condition. In contrast, in total viewing time, the effect of the repetition condition did not reach significance, while both the unrelated and the nonword condition resulted in significantly longer total viewing times. Again, it is likely that the effects in total viewing time are caused by re-reading in second pass. Finally, the gaze durations and single fixation durations on the target word showed a mixed pattern: fixation times in the repetition condition were significantly shorter than those in the control condition, while fixation times in the nonword condition were significantly longer. Again, there was no effect of the preview condition on fixation probability, the probability of making a regression out of the target word, or the landing position on the target word.

INSERT TABLE 4 ABOUT HERE

Post-target word.

The experimental manipulation affected the parafoveal preview that subjects received of the post-target word. As Rayner (1975) showed, the availability of a correct preview of a word (as was the case in our identical control condition) leads to shorter fixation times on that word when it is subsequently fixated. This is known as the *preview benefit effect*. All fixation time measures on the post-target word showed clear evidence of preview benefit effects (see Table 4a for means and Table 4b for the LMM results), with longer fixations in the conditions with a dissimilar preview (i.e., in the repetition, the unrelated, and the nonword conditions) than in the control condition. We observed a similar effect on the probability of making a fixation on the target word and the probability of making a regression out of the target word: when its preview was either an unrelated word or a nonword, readers were more likely to fixate on and, subsequently, make a regression out of the target word than in the control condition. However, there was no difference between the repetition condition and the control condition in terms of fixation and regression probability. There was no effect of the preview condition on landing position on the post-target word.

Discussion

In Experiment 1, we used the gaze-contingent boundary paradigm in order to present readers with the parafoveal repetition of a target word without having to resort to an unusual sentence structure and without making the repetition obvious to the subjects. Indeed, most of the subjects were unaware of the great majority of the display changes. We compared the parafoveal repetition condition with a control condition in which the parafovea contained the actual post-target word and two dissimilar preview conditions in which the parafovea either contained a word that was unrelated to the target and post-target words or a nonword consisting of a random letter string. Our results are quite straightforward: We succeeded in replicating Inhoff et al.'s (2000) finding of shorter fixations on a word when it is repeated in the parafovea. There are two possible explanations for this effect: First, the unusual pattern caused by the parafoveal repetition is detected by the processing system and causes readers to interrupt ongoing processing and make a saccade to the source of the disruption – an attraction effect (Hyönä, 1995) that would result in shorter fixation times on the target word. Second, processing of the target word is actually facilitated by the repetition.

In the former case, we would expect to see evidence of the disruption later, i.e. as longer fixations on the post-target word or as regressions to earlier words in the sentence. However, none of the measures on either the target or the post-target word showed evidence of disruption due to the parafoveal repetition. In contrast, there was some evidence of disruption in the unrelated and nonword preview conditions. On the target word, gaze durations were longer when the parafovea contained a nonword – an orthographic parafoveal-on-foveal effect (Blanchard et al., 1989). The probability of fixating the post-target word and the probability of making a regression out of the post target word were higher in the unrelated and nonword conditions compared to the control condition, which suggests that readers had more difficulties processing the post-target word (and possibly the target word) in those conditions. In short, we were clearly able to find evidence of disruption in the unrelated and nonword conditions, but not in the repetition condition. This result pattern is incompatible with the attraction hypothesis outlined above.

As a consequence, it appears that the facilitation effect we observed was genuine. However, these results do not enable us to make any conclusions as to the processing level on which the facilitation occurred. In particular, the effect could be either orthographic or semantic in nature. Inhoff et al. (2000) found a facilitation effect both when the target word was repeated (*mother's mother*) and when it was followed by a semantically related word (*mother's father*). In order to test this, Experiment 2 included a condition in which the parafoveal information was orthographically related but not identical to the target word and a condition in which the parafoveal word was semantically related to the target word. We also included the identical repetition condition from Experiment 1 as a comparison.

Experiment 2

Method

Subjects.

Forty University of California, San Diego students who had not participated in Experiment 1 or the norming studies participated in this experiment for course credit. All were native speakers of English, had either normal or corrected to normal vision, and were naïve concerning the purpose of the experiment.

Apparatus.

The apparatus was identical to that used in Experiment 1.

Materials.

INSERT FIGURE 2 ABOUT HERE

The experimental sentences used in Experiment 2 were largely similar but not identical to those used in Experiment 1, since it was not possible to find semantically related words of the same length for all of the target words. We used a total of 120 sentences, again manipulating the information available from the parafovea (i.e. the preview of the post-target word) while subjects were fixating the target word. For each target word (e.g. *news)*, we generated an orthographically related nonword neighbor of the target word (*niws*) and selected a semantically related associate (*tale*) along with one of its orthographically unrelated nonword neighbors (*tule*). We asked 17 University of California San Diego students who did not participate in the subsequent experiment to rate the semantic association strength of the target word and the selected semantic associate on a scale from 1 (unrelated) to 7 (strongly related). On average, they rated the semantic association strength as 5.47 (SD .82). Table 5 shows word frequency (from the Corpus of Contemporary American English, Davies, 2011), word length, and mean log bigram frequency (from the *N-Watch* software, Davis, 2005) for pre-target, target, and post-target words.

Procedure.

While subjects were fixating the target word (e.g. *news* in the sentence *Victor read the news once this morning*), the parafoveal preview of the post-target word (see Figure 2) consisted of (1) the actual post-target word (*once*, identical control condition), (2) a repetition of the target word (*news*), (3) an orthographically related nonword neighbor of the target word (*niws*), (4) a semantically related word (*tale*), or (5) an orthographically unrelated nonword (*tule*). Once subjects moved their eyes across the boundary located to the right of the target word, the post-target word preview changed to the actual post-target word (*once*). Note that the control condition and the parafoveal repetition condition were also present in Experiment 1, making Experiment 2 a partial replication of Experiment 1. Again, approximately 33% of the sentences were followed by a two-alternative forced choice comprehension question which subjects answered by pressing the button corresponding to the correct answer on a button box. The mean accuracy was greater than 85% for all subjects

INSERT TABLE 5 ABOUT HERE

Results

We computed the same dependent variables and performed the same LMM analyses for Experiment 2 as for Experiment 1. Again, we used treatment contrasts with the identical control preview condition as the baseline. As in Experiment 1, if any subjects indicated that they had seen more than five display changes their data were excluded from the experiment. This criterion led to the exclusion of 12 out of 52 subjects for a total of 40 subjects included in the analysis. As in Experiment 1, trials with track losses or display changes that completed more than 25 ms after fixation onset as well as trials in which a blink occurred immediately before or during a fixation on the target word were eliminated (2.7% of the data) from the pre-target and target word analyses, except for total viewing time. For the computation of total viewing time on the pre-target and target words and all measures on the post-target word, we used a late display change criterion of 9 ms, which resulted in the elimination of 20% of trials[2]. Additionally, if a fixation was shorter than 80 ms and located within one character space (11 pixels) of another fixation, it was merged into that fixation. All other fixations shorter than 80 ms as well as all fixations longer than 800 ms were deleted (less than 1% of the data). We will now report the results for each of the critical words. Tables 6a, 7a, and 8a show the condition means on the pre-target, target, and post-target while Tables 6b, 7b, and 8b show the results of the LMMs performed.

INSERT TABLE 6 ABOUT HERE

Pre-target word.

There were no significant effects of preview on the pre-target word with the exception of a significant increase in first fixation duration and single fixation duration in the repeated condition compared to the correct control condition. This could be considered a parafoveal-on-foveal effect, possibly caused by the unusual visual repetition pattern in the parafovea. The small effect size suggests that this effect likely does not occur on all trials. Interestingly, there was no such effect in the orthographically related condition, suggesting that complete repetition is necessary for this effect to be observed.

INSERT TABLE 7 ABOUT HERE

Target word.

On the target word, we observed significant facilitation effects for the repeated and the orthographically related conditions compared to the control condition in FFD, SFD, GD, and go-past time. Thus, we successfully replicated the finding in Experiment 1. Furthermore, the orthographically related condition appeared to have a facilitatory effect that was comparable to the repeated condition. There were two other significant effects: In the semantically related and the nonword condition, subjects showed longer total viewing times than in the control condition, while the nonword condition led to a lower fixation probability on the target word. It is not clear to what degree these last effects are interpretable. Perhaps the parafoveal information in the semantically related and the nonword conditions is processed to some degree and causes problems in the late processing stages. No other effects were significant.

INSERT TABLE 8 ABOUT HERE

Post-target word.

As in Experiment 1, we expected the post-target word to show preview benefit effects. This was the case, as all four experimental preview conditions showed significantly longer FFDs, SFDs, GDs, go-past times, and TVTs than the control condition. The preview benefit effect in the semantic condition was numerically smaller than the preview benefit effects in the remaining conditions, at least in early measures, and was only marginally significant in SFD. As there was no semantic overlap between the post-target word (*once*) and the preview word in the semantically related condition (*tale*), this cannot be interpreted as a semantic preview benefit effect. The orthographically related and nonword conditions also resulted in a higher fixation probability on the post-target word. Additionally, the orthographically related, semantically related, and nonword conditions were associated with a higher probability of making regressions out of the target word. No other effects were significant.

Discussion

In Experiment 2, we successfully replicated the parafoveal facilitation effect observed in Experiment 1. Furthermore, while the orthographically related preview condition showed a facilitation effect that was comparable in magnitude to the exact repetition effect, there was no such effect in the semantically related preview condition. This suggests that the repetition preview facilitates processing of the target word due to its orthographic, but not due to its semantic overlap with the target word. This is contrary to Inhoff et al.'s (2000) finding that a semantically related word in the parafovea (*mother's father*) leads to a facilitation effect of similar magnitude as parafoveal repetition of a word (*mother's mother*), compared to an unrelated word (*mother's garden*). One possible explanation for this difference is that there may have been more orthographic overlap between the related word and the target word (e.g. *mother's* and *father*) than between the unrelated word and the target word (e.g. *mother's* and *garden*) in Inhoff et al.'s materials. Another explanation might be that, due to the unusual syntactical structure of the experimental sentences in Inhoff et al.'s experiment, subjects became aware of the manipulation and changed the way they processed the parafoveal word whenever they were fixating a possessive noun.

Somewhat surprisingly, we did not find an interference effect of the nonword condition on fixation times on the target word as in Experiment 1. This may be due to the different character of the nonwords in Experiment 2, which were universally pronounceable and might also be called pseudowords (the mean log bigram frequencies of these words also reflect this difference). It appears that only non-pronounceable nonwords cause measureable orthographic parafoveal-on-foveal effects. In summary, Experiment 2 provided a solid replication of the facilitation effect observed in Experiment 1 and clarified that this effect was not based on semantic, but on orthographic overlap.

General Discussion

In two experiments, we examined whether parafoveal information that overlaps with the information currently available in the fovea or parafovea can facilitate the ongoing processing of the foveal word using the gaze-contingent boundary paradigm (Rayner, 1975). Experiment 1 showed that subjects spent less time fixating a word if an identical repetition of that word was simultaneously presented in the parafovea than if the parafovea contained a random letter string, an unrelated word, or even the actual next word in the sentence. Fixation times on the subsequent word (the actual identity of which was always revealed once a reader made a saccade to fixate it) revealed no evidence of the repetition preview causing disruption or difficulty during further processing beyond the standard preview benefit effect. In Experiment 2, we tested whether the parafoveal facilitation effect was orthographic or semantic in nature. The results indicated that a nonword neighbor of the foveal word provided virtually the same amount of facilitation while a semantically related word provided none. It appears that our results constitute some of the most robust pieces of evidence for a parafoveal-on-foveal effect so far. In the past, most observed parafoveal-on-foveal effects could be explained by mislocated fixations, i.e. readers planning to make a saccade to a parafoveal word, but undershooting due to oculomotor error and fixating the preceding word instead. In this case, fixation times on the actually fixated word would reflect the properties of the parafoveal word that had been the target of the saccade (for a review, see Schotter et al., 2012). The parafoveal-foveal overlap effect we observed cannot be explained by mislocated fixations, since it not only reflects the properties of the parafoveal word but the overlap between the parafoveal word and the foveal word. This suggests that parafoveal-foveal overlap affects ongoing processing on the currently fixated word.

In summary, parafoveal information can indeed influence ongoing processing, but only if it shares orthographic features with the foveal word. This is a surprising finding that no current models of reading can account for: Neither E-Z Reader nor SWIFT have an explicit mechanism that would allow orthographic parafoveal information to influence fixation times on the currently fixated word[3].The model that is closest to the framework we will propose below is Glenmore. However, it too lacks connections between parafoveal letter units and the currently fixated word node.

There is the possibility that parafoveal-foveal overlap might reduce competition between the foveal and the parafoveal word and therefore make the foveal word easier to identify, which in turn may influence the fixate center and cause a change in fixation time. Simulations using the computational implementation of the Glenmore model might show whether this is actually the case.

How might the facilitatory effect of parafoveal overlap be explained in general terms? We propose a simple explanation which can be incorporated into a variety of distributed word recognition models and that is consistent with serial attention shift (SAS) accounts of eye movement control in reading (such as E-Z Reader). For simplicity, we will describe our approach only in terms of the Interactive Activation (IA) model, but this does not mean that it is specific for that model; it could just as easily be incorporated into other models such as the Parallel Distributed Processing (PDP) model (Seidenberg & McClelland, 1989) or the Dual Route Cascaded (DRC) model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001).

INSERT FIGURE 3 ABOUT HERE

Letter detectors receive both parafoveal and foveal input

In essence, our approach assumes that parafoveal and foveal information interact on the letter level. In this scenario, there is a single set of letter detectors representing the letters of the currently processed word. These letter nodes receive input from both foveal and parafoveal feature detectors, with visual attention giving priority to the word that is currently being processed by enhancing the activation level of its features (Figure 3). This has some parallels with the pre-attentive visual stage in E-Z Reader, but differs from it in that it assumes that attention is already deployed elsewhere and in that it makes concrete predictions as to how this pre-attentive information affects ongoing processing. As in E-Z Reader, we assume that lexical processing occurs serially and requires attention. Consequently, once processing of the foveal word has finished, attention shifts to the parafoveal word (with a saccade to that word being programmed in parallel).Even though we only make a few additional assumptions beyond those of the E-Z Reader model, these assumptions enable us to make some interesting predictions: Due to the connections between parafoveal feature detectors and letter detectors, the letters of the upcoming word should be able to affect processing on the current word. In particular, if the upcoming word shares letters with the current word, this should facilitate processing of the current word – in effect, the model can explain Inhoff et al.'s (2000) and our findings.

On the other hand, if the upcoming word has letters that are different from the letters of the foveal word, information about these letters may compete with the foveal letters and slow down the recognition of the foveal word. This alone does not explain orthographic parafoveal-on-foveal effects, since the foveal word and the successor word are very unlikely to be identical in any situation. However, it is possible that unusual letter strings disrupt foveal processing more strongly than the letter strings in a parafoveal word (or a word-like nonword). This might be due to an intermediate layer between letters and words that encodes bigrams or possibly syllables. In any case, there is likely to be more overlap, on average, between the foveal word and a parafoveal word than between the foveal word and a random letter string. This difference alone might be enough to drive parafoveal-on-foveal effects, which are usually quite small in size. Importantly, this prediction can easily be tested, using existing reading data, by using a linear regression model with fixation time on a word as the dependent variable and the amount of overlap between any word and its successor as a predictor while controlling for known variables such as word length and word frequency. If we assume (as all current models of eye movement control in reading do) that processing of letters is slower the larger the eccentricity at which they are located, this framework suggests that parafoveal facilitation either does not occur or is extremely weak when the repetition preview is not available while the target word is being fixated.

In summary, the framework we have described has the potential to explain orthographic parafoveal-on-foveal effects on fixation times as described above. Importantly, we showed that our explanation is perfectly compatible with the notion of serial attention shifts as implemented in the E-Z Reader model and does not require multiple words being processed in parallel. Future research will be able to test whether this effect is limited to the orthographic level or whether it might also occur at the phonological and morpheme levels.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi:10.1016/j.jml.2007.12.005

Bates, D., Maechler, M., & Dai, B. (2009). lme4: Linear mixed-effect models using S4 classes (R package Ver. 0.999375-28). *Vienna, Austria: R Foundation for Statistical Computing*.

Blanchard, H. E., Pollatsek, A., & Rayner, K. (1989). The acquisition of parafoveal word information in reading. *Perception & Psychophysics*, *46*, 85 – 94.

Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. *Cognitive Processes in Eye Guidance*, 53–77.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.

Davies, M. (2011). Word frequency data from the Corpus of Contemporary American English (COCA). *Retrieved from Wordfrequency.info on December 19, 2011*.

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*, 65–70.

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.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge, UK: Cambridge University 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.

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 nonfixated words influence the processing of fixated words during reading? *European Journal of Cognitive Psychology*, *16*, 104–127.

Inhoff, A. W., Radach, R., Starr, M., & Greenberg, S. (2000). Allocation of visuospatial 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: North-Holland/Elsevier.

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. *The Quarterly Journal of Experimental Psychology Section A*, *53*, 429–455.

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

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.

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

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

R Development Core Team, R. (2011). R: a language and environment for statistical computing. *R Foundation for Statistical Computing*.

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

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

Rayner, K. (2009). The Thirty-Fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*, 1457–1506.

Rayner, K., & Clifton, C. (2009). Language Processing in Reading and Speech Perception is Fast and Incremental: Implications for Event Related Potential Research. *Biological Psychology*, *80*, 4–9. doi:10.1016/j.biopsycho.2008.05.002

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., Li, X., & Pollatsek, A. (2007). Extending the E-Z reader model of eye movement control to Chinese readers. *Cognitive Science*, *31*, 1021–1033.

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

Reichle, E. D., 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., 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.

Reilly, R. G., & Radach, R. (2003). Glenmore: An interactive activation model of eye movement control in reading. *ICONIP'02. Proceedings of the 9th International Conference on Neural Information Processing, 2002.* (Vol. 3, pp. 1194–1200).

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.

Risse, S., Engbert, R., & Kliegl, R. (2008). Eye-movement control in reading: Experimental and corpus-analytic challenges for a computational model. In K. Rayner, D. Shen, X. Bai, and G. Yan (Eds), *Cognitive and Cultural Influences on Eye Movements* (Tianjin, China., pp. 65–91). Tianjin People's Publishing House.

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

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

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

White, S. J., & Liversedge, S. P. (2006). Linguistic and nonlinguistic influences on the eyes' landing positions during reading. *The 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 reexamination. *Psychonomic Bulletin & Review*, *12*, 891–896. doi:10.3758/BF03196782

White, S., & Liversedge, S. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology*, *16*, 52–78.

Footnotes

1. We also performed the same analyses with a stricter criterion (3 ms) which led to an exclusion of 56.5% of trials. The pattern of results in the lenient analysis did not significantly differ from the pattern observed in the more restrictive 3 ms analysis.

2. As in Experiment 1, we also performed the same analysis with a stricter criterion (3 ms) which led to an exclusion of 52% of trials. Just like in Experiment 1, the pattern of results in the lenient analysis did not significantly differ from the pattern observed in the more restrictive 3 ms analysis.

3. SWIFT can potentially account for parafoveal-on-foveal effects in gaze duration, but it cannot explain the effects we observed in first fixation duration (Risse, Engbert, & Kliegl, 2008). Additionally, there is no mechanism in SWIFT that can combine parafoveal and foveal information.

Acknowledgments

This research was supported by grant HD26765 from the National Institute of Health. We thank Barbara Juhasz and one anonymous reviewer for helpful comments and suggestions. Correspondence should be sent to Bernhard Angele (bangele@ucsd.edu).

Table 1: Preview word properties in Experiment 1. Word frequencies were obtained from the Corpus of Contemporary American English (Davies, 2011); mean log bigram frequency values were obtained from the *N-Watch* software (Davis, 2005).

Condition	**Word Frequency**	**Mean Log Bigram Frequency**
Identical control	79.26 (180.94)	2.86 (0.44)
Repeated	112.17 (234.19)	2.84 (0.46)
Unrelated	56.99 (235.93)	2.67 (0.60)
Nonword	--	1.47 (0.68)

Standard deviations are in parentheses.

Table 2a: Condition means on the pre-target word.

	Fixation time measures					Probabilities		Fixation location measures
	First fixation duration	Single fixation duration	Gaze duration	Go-past time	Total viewing time	Fixation probability	Probability of regressions out	Landing position
correct	216 (73.6)	219 (75.4)	234 (92.3)	257 (131)	257 (129)	0.553 (0.497)	0.0404 (0.197)	1.87 (1.58)
repeated	212 (70.4)	213 (69.3)	230 (91.7)	262 (154)	263 (141)	0.563 (0.496)	0.0433 (0.204)	1.86 (1.59)
unrelated	212 (69.6)	214 (69.9)	239 (102)	270 (162)	278 (156)	0.54 (0.499)	0.0356 (0.185)	1.88 (1.63)
nonword	213 (67.6)	214 (68)	237 (97.2)	266 (149)	275 (164)	0.575 (0.495)	0.039 (0.194)	1.88 (1.61)

Standard deviations are in parentheses.

Table 2b: LMM analyses on the pre-target word. Each column represents a model.

		First fixation duration			Single fixation duration			Gaze duration			Go-past time			Total viewing time			Fixation probability			Probability of regressions out			Landing position		
		b	SE	t	b	SE	t	b	SE	t	b	SE	t	B	SE	t	b	SE	z	b	SE	z	b	SE	t
	(Intercept)	214.28	4.62	**46.38**	217.57	4.87	**44.67**	227.80	6.28	**36.28**	250.88	9.09	**27.60**	246.46	10.62	**23.21**	0.38	0.16	**2.42**	-3.47	0.18	**-18.91**	1.65	0.09	**18.45**
Preview	Repeated*	-4.21	3.51	-1.20	-5.47	3.71	-1.48	-4.51	4.68	-0.96	3.11	7.46	0.42	4.02	8.48	0.47	0.07	0.09	0.75	0.08	0.21	0.37	-0.03	0.07	-0.39
	Unrelated*	-3.98	3.55	-1.12	-4.21	3.79	-1.11	3.96	4.74	0.84	13.18	7.55	1.74	18.57	8.60	**2.16**	-0.07	0.09	-0.76	-0.15	0.22	-0.70	-0.03	0.08	-0.43
	Nonword*	-2.96	3.50	-0.85	-4.55	3.73	-1.22	2.31	4.66	0.50	8.26	7.44	1.11	18.94	8.68	**2.18**	0.12	0.09	1.27	-0.04	0.21	-0.19	0.00	0.07	0.00
Random effects	Item	227.10	15.07	NA	360.54	18.99	NA	839.38	28.97	NA	1667.10	40.83	NA	2759.70	52.53	NA	1.81	1.35	NA	0.37	0.61	NA	0.62	0.79	NA
	Subject	528.95	23.00	NA	555.55	23.57	NA	863.73	29.39	NA	1631.00	40.38	NA	2089.60	45.71	NA	0.28	0.53	NA	0.32	0.56	NA	0.01	0.12	NA
	Residual	4201.98	64.82	NA	4094.26	63.99	NA	7431.01	86.20	NA	18951.50	137.66	NA	17124.00	130.86	NA	NA	NA	NA	NA	NA	NA	1.89	1.37	NA

*Compared to the correct control condition as the baseline.
b: Regression coefficient, SE: standard error, t: test statistic (b/SE). Cell s with |t| >= 1.96 are marked in bold.

Table 3a: Condition means on the target word

	Fixation time measures					Probabilities		Fixation location measures
	First fixation duration	**Single fixation duration**	**Gaze duration**	**Go-past time**	**Total viewing time**	**Fixation probability**	**Probability of regressions out**	**Landing position**
correct	230 (76.1)	233 (75.6)	262 (107)	304 (176)	307 (151)	0.834 (0.372)	0.0825 (0.275)	2.06 (1.45)
repeated	222 (70.1)	226 (70.7)	250 (100)	278 (143)	296 (148)	0.86 (0.347)	0.0617 (0.241)	2.12 (1.48)
unrelated	232 (81.7)	233 (79.2)	267 (116)	310 (178)	347 (193)	0.842 (0.365)	0.0808 (0.273)	2.08 (1.44)
nonword	236 (77.9)	241 (77.8)	278 (119)	317 (175)	351 (190)	0.835 (0.371)	0.0771 (0.267)	2.13 (1.47)

Standard deviations are in parentheses.

Table 3b: LMM analyses on the target word. Each column represents a model.

		First fixation duration			Single fixation duration			Gaze duration			Go-past time			Total viewing time			Fixation probability			Probability of regressions out			Landing position		
		b	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**B**	**SE**	**t**	**b**	**SE**	**z**	**b**	**SE**	**z**	**b**	**SE**	**t**
	(Intercept)	229.56	4.52	**50.82**	233.56	5.05	**46.28**	259.54	7.10	**36.54**	300.71	10.06	**29.90**	305.98	11.31	**27.05**	1.98	0.18	**10.80**	-2.68	0.15	**-17.36**	2.07	0.06	**32.01**
Preview	Repeated*	-7.94	3.13	**-2.54**	-7.22	3.36	**-2.15**	-11.65	4.47	**-2.61**	-25.51	6.88	**-3.71**	-14.44	8.15	-1.77	0.26	0.12	**2.10**	-0.33	0.16	**-2.07**	0.06	0.06	0.91
	Unrelated*	1.14	3.15	0.36	-0.85	3.45	-0.25	4.45	4.51	0.99	4.86	6.93	0.70	35.07	8.29	**4.23**	0.08	0.12	0.65	-0.03	0.15	-0.18	0.02	0.06	0.32
	Nonword*	5.08	3.16	1.61	7.70	3.47	**2.22**	15.58	4.52	**3.45**	12.75	6.95	1.83	43.80	8.41	**5.21**	-0.01	0.12	-0.06	-0.08	0.15	-0.54	0.07	0.06	1.09
Random effects	Item	265.85	16.30	NA	316.11	17.78	NA	562.06	23.71	NA	1005.90	31.72	NA	1581.30	39.77	NA	0.29	0.53	NA	0.29	0.54	NA	0.11	0.33	NA
	Subject	534.29	23.11	NA	684.65	26.17	NA	1436.75	37.90	NA	2768.90	52.62	NA	3232.20	56.85	NA	0.95	0.98	NA	0.39	0.62	NA	0.06	0.24	NA
	Residual	5113.32	71.51	NA	4804.35	69.31	NA	10398.65	101.97	NA	24735.40	157.28	NA	24732.50	157.27	NA	NA	NA	NA	NA	NA	NA	1.98	1.41	NA

*Compared to the correct control condition as the baseline.
b: Regression coefficient, SE: standard error, t: test statistic (b/SE). Cell s with |t| >= 1.96 are marked in bold.

Table 4a: Condition means on the post-target word

	Fixation time measures					Probabilities		Fixation location measures
	First fixation duration	Single fixation duration	Gaze duration	Go-past time	Total viewing time	Fixation probability	Probability of regressions out	Landing position
correct	250 (92.3)	256 (92.7)	279 (115)	309 (159)	314 (152)	0.838 (0.369)	0.0638 (0.245)	2.8 (1.27)
repeated	287 (111)	303 (109)	325 (127)	369 (189)	381 (172)	0.858 (0.349)	0.084 (0.278)	2.85 (1.23)
unrelated	283 (109)	297 (107)	318 (126)	389 (203)	392 (196)	0.873 (0.334)	0.133 (0.34)	2.82 (1.23)
nonword	290 (115)	302 (108)	322 (125)	400 (219)	398 (205)	0.904 (0.295)	0.15 (0.357)	2.75 (1.24)

Standard deviations are in parentheses.

Table 4b: LMM analyses on the post-target word. Each column represents a model.

		First fixation duration			Single fixation duration			Gaze duration			Go-past time			Total viewing time			Fixation probability			Probability of regressions out			Landing position		
		b	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**B**	**SE**	**t**	**b**	**SE**	**z**	**b**	**SE**	**z**	**b**	**SE**	**t**
	(Intercept)	248.11	7.05	**35.20**	256.05	7.73	**33.13**	277.15	8.30	**33.41**	308.72	12.29	**25.12**	311.32	12.20	**25.51**	1.94	0.19	**10.38**	-2.89	0.18	**-16.19**	2.83	0.08	**35.41**
Preview	Repeated*	38.00	5.42	**7.02**	47.30	5.69	**8.31**	45.54	6.18	**7.37**	55.44	9.78	**5.67**	65.72	8.85	**7.43**	0.18	0.15	1.23	0.28	0.20	1.40	0.04	0.06	0.62
	Unrelated*	34.73	5.45	**6.37**	40.73	5.80	**7.03**	40.28	6.22	**6.47**	77.93	9.86	**7.91**	78.23	8.88	**8.81**	0.34	0.15	**2.22**	0.84	0.18	**4.54**	-0.01	0.06	-0.21
	Nonword*	42.05	5.49	**7.67**	46.41	5.83	**7.96**	43.57	6.27	**6.95**	89.99	9.90	**9.09**	87.47	9.00	**9.72**	0.67	0.17	**4.04**	1.01	0.18	**5.48**	-0.07	0.06	-1.06
Random effects	Item	337.69	18.38	NA	759.78	27.56	NA	780.82	27.94	NA	2535.70	50.36	NA	2163.00	46.51	NA	0.16	0.39	NA	0.25	0.50	NA	0.06	0.25	NA
	Subject	1261.51	35.52	NA	1484.37	38.53	NA	1706.15	41.31	NA	3237.40	56.90	NA	3626.10	60.22	NA	0.89	0.94	NA	0.27	0.52	NA	0.15	0.39	NA
	Residual	9964.29	99.82	NA	8821.62	93.92	NA	12857.29	113.39	NA	32292.60	179.70	NA	28038.30	167.45	NA	NA	NA	NA	NA	NA	NA	1.33	1.15	NA

*Compared to the correct control condition as the baseline.
b: Regression coefficient, SE: standard error, t: test statistic (b/SE). Cell s with |t| >= 1.96 are marked in bold.

Table 5: Properties of the target region in Experiment 2. Word frequencies were obtained from the Corpus of Contemporary American English (Davies, 2011); mean log bigram frequency values were obtained from the *N-Watch* software (Davis, 2005).

Condition	Mean word Frequency	Mean Log Bigram Frequency
Identical control	79.26 (180.94)	2.86 (0.44)
Repeated	112.17 (234.19)	2.84 (0.46)
Orthographic	0.01 (0.03)	2.41 (0.56)
Semantic	108.62 (201.56)	2.83 (0.50)
Nonword	0.78 (8.20)	2.41 (0.55)

Standard deviations are in parentheses.

Table 6a: Condition means on the pre-target word in Experiment 2

	Fixation time measures					Probabilities		Fixation location measures
	First fixation duration	**Single fixation duration**	**Gaze duration**	**Go-past time**	**Total viewing time**	**Fixation probability**	**Probability of regressions out**	**Landing position**
Identical control	198 (53.8)	198 (52.8)	211 (70.7)	247 (144)	251 (146)	0.516 (0.5)	0.0458 (0.209)	3.11 (2.6)
Repeated	206 (64)	207 (65.4)	222 (85.5)	252 (135)	261 (144)	0.519 (0.5)	0.0511 (0.22)	3.56 (3.06)
Orthographic	200 (63.8)	202 (63.8)	218 (86.8)	244 (124)	268 (165)	0.493 (0.5)	0.0434 (0.204)	3.57 (2.59)
Semantic	202 (61.5)	203 (62.7)	215 (80)	246 (135)	264 (157)	0.516 (0.5)	0.0502 (0.219)	4 (2.65)
Nonword	204 (63.2)	205 (62.1)	219 (80.7)	243 (131)	269 (152)	0.513 (0.5)	0.0404 (0.197)	3.78 (2.83)

Standard deviations are in parentheses.

Table 6b: LMM analyses on the pre-target word in Experiment 2. Each column represents a model.

		First fixation duration			**Single fixation duration**			**Gaze duration**			**Go-past time**			**Total viewing time**			**Fixation probability**			**Probability of regressions out**			**Landing position**		
		b	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**B**	**SE**	**t**	**b**	**SE**	**z**	**b**	**SE**	**z**	**b**	**SE**	**t**
	(Intercept)	196.11	4.78	**41.07**	196.56	5.05	**38.92**	206.27	6.12	**33.70**	243.90	9.19	**26.53**	241.90	11.87	**20.38**	0.16	0.15	1.03	-3.32	0.20	**-16.27**	2.55	0.34	**7.42**
Preview	Repeated*	7.70	3.60	**2.14**	8.83	3.78	**2.34**	9.12	4.69	1.94	4.12	8.14	0.51	4.93	9.33	0.53	0.03	0.11	0.28	0.11	0.23	0.48	0.36	0.35	1.03
	Orthographic*	2.33	3.66	0.64	4.37	3.85	1.13	6.01	4.77	1.26	-4.30	8.27	-0.52	16.61	9.37	1.77	-0.12	0.11	-1.10	-0.05	0.23	-0.22	-0.14	0.36	-0.38
	Semantic*	4.22	3.62	1.17	5.59	3.77	1.48	3.46	4.71	0.74	-1.94	8.17	-0.24	11.87	9.37	1.27	0.01	0.11	0.05	0.11	0.23	0.50	0.13	0.39	0.33
	Nonword*	5.61	3.63	1.54	6.76	3.80	1.78	6.18	4.74	1.30	-5.37	8.21	-0.65	13.45	9.40	1.43	-0.04	0.11	-0.37	-0.14	0.24	-0.57	0.01	0.37	0.01
Random effects	Item	195.31	13.97	NA	274.86	16.58	NA	597.68	24.45	NA	1190.20	34.50	NA	1631.40	40.39	NA	1.58	1.26	NA	0.18	0.42	NA	3.34	1.83	NA
	Subject	577.78	24.04	NA	631.10	25.12	NA	836.97	28.93	NA	1610.50	40.13	NA	3279.80	57.27	NA	0.19	0.43	NA	0.49	0.70	NA	0.45	0.67	NA
	Residual	2964.39	54.45	NA	2865.05	53.53	NA	5014.17	70.81	NA	15173.70	123.18	NA	18233.50	135.03	NA	NA	NA	NA	NA	NA	NA	2.47	1.57	NA

*Compared to the correct control condition as the baseline.
b: Regression coefficient, SE: standard error, t: test statistic (b/SE). Cell s with |t| >= 1.96 are marked in bold.

Table 7a: Condition means on the target word in Experiment 2

	Fixation time measures					Probabilities		Fixation location measures
	First fixation duration	**Single fixation duration**	**Gaze duration**	**Go-past time**	**Total viewing time**	**Fixation probability**	**Probability of regressions out**	**Landing position**
Identical control	223 (68.9)	224 (68.6)	241 (89.7)	282 (165)	307 (169)	0.836 (0.371)	0.0894 (0.285)	3.26 (1.6)
Repeated	214 (63.7)	215 (64)	229 (80.1)	264 (156)	295 (169)	0.804 (0.398)	0.0699 (0.255)	2.9 (1.7)
Orthographic	214 (60.8)	214 (59.7)	232 (85.5)	259 (132)	311 (184)	0.816 (0.387)	0.0756 (0.265)	2.86 (1.81)
Semantic	223 (75.7)	225 (72.8)	243 (97.9)	281 (149)	338 (186)	0.806 (0.396)	0.0926 (0.29)	3.25 (1.53)
Nonword	222 (72.5)	225 (72.7)	243 (95.3)	288 (160)	328 (179)	0.793 (0.405)	0.109 (0.312)	3.02 (1.66)

Standard deviations are in parentheses.

Table 7b: LMM analyses on the target word in Experiment 2. Each column represents a model.

		First fixation duration			**Single fixation duration**			**Gaze duration**			**Go-past time**			**Total viewing time**			**Fixation probability**			**Probability of regressions out**			**Landing position**		
		b	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**B**	**SE**	**t**	**b**	**SE**	**z**	**b**	**SE**	**z**	**b**	**SE**	**t**
	(Intercept)	222.27	4.48	**49.58**	224.09	4.70	**47.68**	239.97	5.78	**41.52**	280.77	9.43	**29.78**	305.44	13.13	**23.26**	1.92	0.17	**11.13**	-2.65	0.18	**-15.03**	3.22	0.18	**17.63**
Preview	Repeated*	-8.69	3.32	**-2.62**	-7.84	3.50	**-2.24**	-12.33	4.35	**-2.83**	-18.65	7.44	**-2.51**	-14.57	8.69	-1.68	-0.24	0.13	-1.87	-0.30	0.18	-1.67	-0.31	0.23	-1.36
	Orthographic*	-8.21	3.31	**-2.48**	-9.06	3.49	**-2.60**	-9.67	4.33	**-2.23**	-23.12	7.42	**-3.12**	5.41	8.65	0.63	-0.16	0.13	-1.23	-0.21	0.18	-1.16	-0.37	0.23	-1.61
	Semantic*	-0.07	3.32	-0.02	0.59	3.53	0.17	1.02	4.35	0.24	-2.38	7.44	-0.32	34.94	8.72	**4.01**	-0.22	0.13	-1.69	0.07	0.17	0.40	0.03	0.22	0.13
	Nonword*	-0.26	3.34	-0.08	1.16	3.56	0.33	2.03	4.37	0.46	6.40	7.48	0.86	20.13	8.79	**2.29**	-0.32	0.13	**-2.46**	0.24	0.17	1.45	-0.30	0.22	-1.41
Random effects	Item	134.57	11.60	NA	169.05	13.00	NA	381.68	19.54	NA	1143.10	33.81	NA	1697.40	41.20	NA	0.23	0.48	NA	0.38	0.62	NA	0.12	0.35	NA
	Subject	541.40	23.27	NA	582.67	24.14	NA	835.59	28.91	NA	2077.80	45.58	NA	4813.20	69.38	NA	0.74	0.86	NA	0.50	0.70	NA	0.28	0.53	NA
	Residual	4031.48	63.49	NA	3857.88	62.11	NA	6894.91	83.04	NA	20206.40	142.15	NA	25089.70	158.40	NA	NA	NA	NA	NA	NA	NA	2.38	1.54	NA

*Compared to the correct control condition as the baseline.
b: Regression coefficient, SE: standard error, t: test statistic (b/SE). Cell s with |t| >= 1.96 are marked in bold.

Table 8a: Condition means on the post-target word in Experiment 2

	Fixation time measures					Probabilities		
	First fixation duration	**Single fixation duration**	**Gaze duration**	**Go-past time**	**Total viewing time**	**Fixation probability**	**Probability of regressions out**	**Landing position**
Identical control	234 (75.1)	240 (77)	260 (97.6)	312 (176)	330 (177)	0.806 (0.396)	0.101 (0.302)	3 (1.51)
Repeated	260 (89.1)	269 (87.7)	296 (115)	345 (189)	359 (193)	0.824 (0.381)	0.0976 (0.297)	3.13 (1.55)
Orthographic	253 (85)	263 (83.5)	286 (112)	348 (189)	362 (192)	0.854 (0.353)	0.134 (0.341)	2.67 (1.76)
Semantic	245 (87.3)	247 (86.8)	272 (112)	349 (216)	377 (207)	0.818 (0.386)	0.147 (0.354)	2.96 (1.62)
Nonword	248 (85.9)	254 (82)	279 (113)	360 (212)	365 (189)	0.869 (0.338)	0.173 (0.378)	2.84 (1.54)

Standard deviations are in parentheses.

Table 8b: LMM analyses on the post-target word in Experiment 2. Each column represents a model.

		First fixation duration			Single fixation duration			Gaze duration			Go-past time			Total viewing time			Fixation probability			Probability of regressions out			Landing position		
		b	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**b**	**SE**	**t**	**B**	**SE**	**t**	**b**	**SE**	**z**	**b**	**SE**	**z**	**b**	**SE**	**t**
	(Intercept)	232.67	5.40	**43.13**	238.05	6.01	**39.63**	257.61	6.77	**38.05**	308.58	12.38	**24.93**	325.37	14.80	**21.98**	1.74	0.19	**9.22**	-2.47	0.18	**-14.08**	2.90	0.18	**16.37**
Preview	Repeated*	25.22	4.59	**5.49**	27.75	4.98	**5.57**	34.21	6.01	**5.69**	28.50	10.57	**2.70**	28.07	9.31	**3.01**	0.11	0.14	0.75	-0.09	0.18	-0.51	0.13	0.19	0.70
	Orthographic*	17.90	4.53	**3.95**	22.30	4.92	**4.53**	26.93	5.94	**4.54**	35.06	10.44	**3.36**	32.98	9.24	**3.57**	0.41	0.15	**2.76**	0.33	0.17	**1.97**	-0.29	0.19	-1.55
	Semantic*	11.16	4.61	**2.42**	9.20	5.04	1.82	12.92	6.04	**2.14**	38.77	10.61	**3.65**	49.90	9.35	**5.34**	0.08	0.14	0.57	0.47	0.17	**2.80**	-0.11	0.19	-0.56
	Nonword*	15.14	4.58	**3.31**	16.74	5.01	**3.34**	19.28	6.00	**3.21**	47.92	10.54	**4.54**	37.48	9.36	**4.00**	0.53	0.15	**3.46**	0.66	0.16	**4.03**	-0.13	0.18	-0.72
Random effects	Item	215.66	14.69	NA	338.55	18.40	NA	487.17	22.07	NA	2788.40	52.80	NA	2327.30	48.24	NA	0.24	0.49	NA	0.29	0.54	NA	0.22	0.47	NA
	Subject	657.37	25.64	NA	828.99	28.79	NA	925.35	30.42	NA	2889.60	53.76	NA	6180.60	78.62	NA	0.91	0.95	NA	0.47	0.69	NA	0.34	0.58	NA
	Residual	6323.34	79.52	NA	5876.08	76.66	NA	10794.02	103.89	NA	33237.40	182.31	NA	28470.60	168.73	NA	NA	NA	NA	NA	NA	NA	2.07	1.44	NA

*Compared to the correct control condition as the baseline.
b: Regression coefficient, SE: standard error, t: test statistic (b/SE). Cell s with |t| >= 1.96 are marked in bold.

Figure captions

1. Condition examples and display change procedure in Experiment 1

2. Condition examples and display change procedure in Experiment 2

3. A simple model framework capable of explaining parafoveal repetition effects and orthographic parafoveal-on-foveal effects

Figure 1

Identical control preview	`Victor read the news	once this morning.`
Repetition preview	`Victor read the news	news this morning.`
Unrelated preview	`Victor read the news	warm this morning.`
Nonword preview	`Victor read the news	rzmc this morning.`

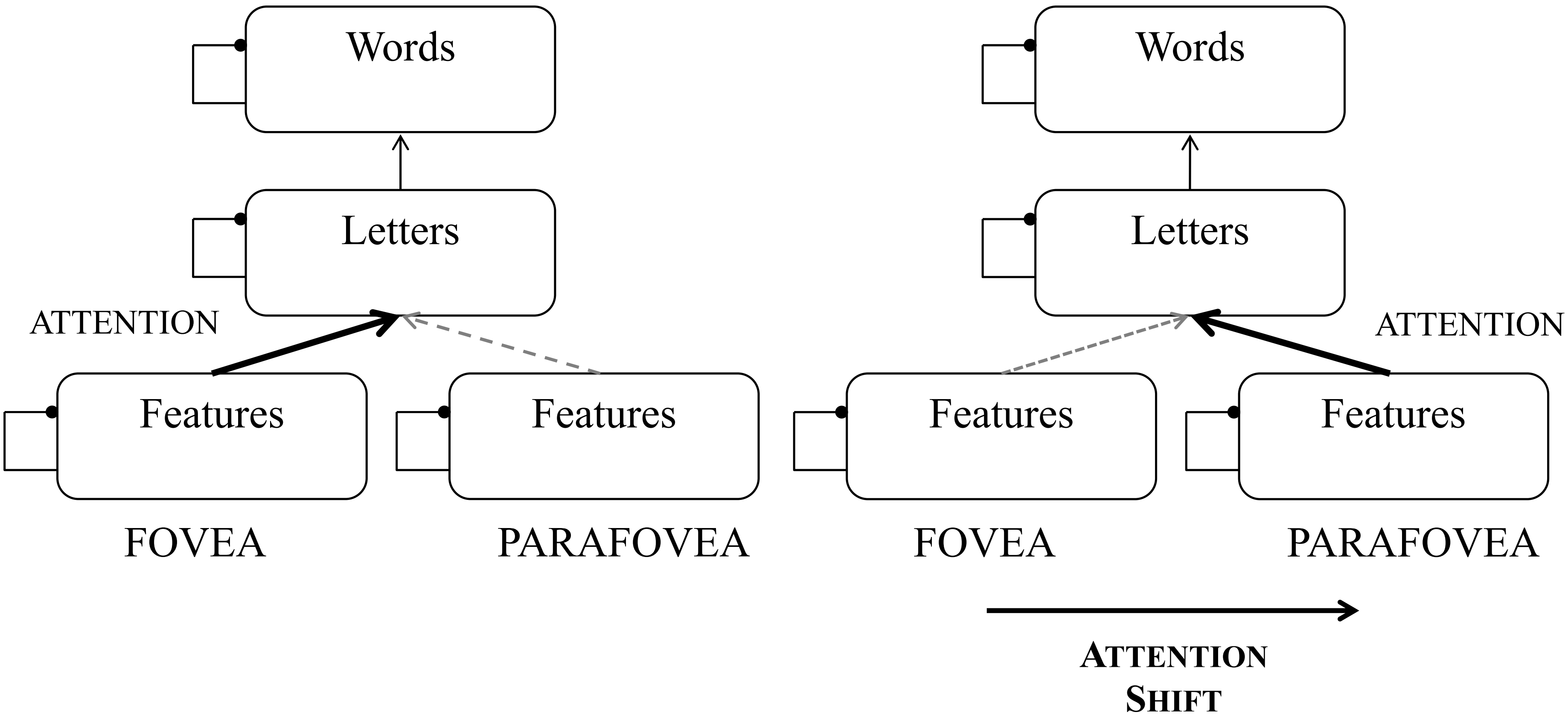
The dashed line represents the invisible boundary. After it was crossed, the display always showed the identical control stimulus.

Figure 2

Identical control preview	Victor read the news once this morning.
Repetition preview	Victor read the news news this morning.
Orthographically related preview	Victor read the news niws this morning.
Semantically related preview	Victor read the news tale this morning.
Nonword preview	Victor read the news tule this morning.

The dashed line represents the invisible boundary. After it was crossed, the display always showed the identical control stimulus.

Figure 3



Words
Letters
ATTENTION
Features
Features
FOVEA
PARAFOVEA
Words
Letters
ATTENTION
Features
Features
FOVEA
PARAFOVEA
ATTENTION
SHIFT