# Eye movements during reading of randomly shuffled text

**Citation for published version:**
Schad, DJ, Nuthmann, A & Engbert, R 2010, 'Eye movements during reading of randomly shuffled text' Vision Research, vol. 50, no. 23, pp. 2600-2616. DOI: 10.1016/j.visres.2010.08.005

**Digital Object Identifier (DOI):**
10.1016/j.visres.2010.08.005



**Published In:**
Vision Research

Vision Research 50 (2010) 2600–2616

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

# Eye movements during reading of randomly shuffled text

Daniel J. Schad [a,*], Antje Nuthmann [b], Ralf Engbert [a]

[a] Department of Psychology, University of Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany
[b] Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK

ARTICLE INFO

*Article history:*
Received 6 May 2010
Received in revised form 30 July 2010

*Keywords:*
Reading
Eye movements
Linear mixed-effect model
Perceptual span
Foveal load hypothesis
Parafoveal-on-foveal effects

ABSTRACT

In research on eye-movement control during reading, the importance of cognitive processes related to language comprehension relative to visuomotor aspects of saccade generation is the topic of an ongoing debate. Here we investigate various eye-movement measures during reading of randomly shuffled meaningless text as compared to normal meaningful text. To ensure processing of the material, readers were occasionally probed for words occurring in normal or shuffled text. For reading of shuffled text we observed longer fixation times, less word skippings, and more refixations than in normal reading. Shuffled-text reading further differed from normal reading in that low-frequency words were not overall fixated longer than high-frequency words. However, the frequency effect was present on long words, but was reversed for short words. Also, consistent with our prior research we found distinct experimental effects of spatially distributed processing over several words at a time, indicating how lexical word processing affected eye movements. Based on analyses of statistical linear mixed-effect models we argue that the results are compatible with the hypothesis that the perceptual span is more strongly modulated by foveal load in the shuffled reading task than in normal reading. Results are discussed in the context of computational models of reading.



## 1. Introduction

Reading represents a very complex task because some of the key cognitive systems (e.g., vision, attention, word recognition, memory, oculomotor control, higher-level language comprehension) must interact to move the eyes across the text. Measurement of eye movements represents a powerful approach to investigate the cognitive subsystems involved in reading as eye movements provide a sensitive online-measure for these processes (Rayner, 1998, 2009). One of the most important problems in current research on the control of eye movements concerns the relative importance of low-level visuomotor processes vs. higher-level cognition related to language processing (Starr & Rayner, 2001). This research problem extends to other aspects of *active vision*, where eye movements are needed for visual information uptake (Liversedge & Findlay, 2000).

Computational models of reading implement theories about how different cognitive processes act in concert to control the movements of the eyes (for an overview of current models, see the 2006 special issue of *Cognitive Systems Research*). It is undisputed that low-level processes like visual perception and oculomotor control affect eye movements during reading. *Primary oculomotor control* models (POC) focus on such low-level processes and ignore direct cognitive influences on eye movements (e.g., Reilly & O'Regan, 1998). *Cognitive models*, to the contrary, assume that higher-level cognition related to language processing plays an important part in controlling the eyes (e.g., E-Z Reader: Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006; Reichle, Warren, & McConnell, 2009; SWIFT: Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005).

Up to now, computational models have mainly considered two kinds of cognitive influences on eye movements. The first one is the lexical processing of words, i.e., the type of processing that is needed to get access to a word's entry in the mental lexicon (e.g., Engbert et al., 2002; Morrison, 1984; Reichle et al., 1998; Reilly & Radach, 2006). The second cognitive influence concerns the predictions that readers make about upcoming words in a text (e.g., Engbert et al., 2002; Reichle et al., 1998). Recently, a first attempt has been made to also include some effects of higher-level language processing in a computational model of reading (Reichle, Warren et al., 2009).

Two general strategies have been used to test hypotheses about how higher- and lower-level factors influence eye movements. First, processes can be tied to the influence of certain variables that modulate these effects. For example, word length is regarded as a low-level variable affecting visual processing. Typically, readers look longer at long words than at short words (e.g., Just & Carpenter, 1980; Kliegl, Nuthmann, & Engbert, 2006; Rayner, Sereno, & Raney, 1996). Effects of word frequency and word predictability,

* Corresponding author. Fax: +49 331 977 2793.
*E-mail address:* Daniel.Schad@uni-potsdam.de (D.J. Schad).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2010.08.005

to the contrary, are thought to result from higher-level cognitive influences on eye movements. Low-frequency words are fixated longer than high-frequency words (e.g., Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kliegl et al., 2006; Rayner & Duffy, 1986). This is mainly because word frequency affects lexical processing, i.e., it takes longer to recognize words that do not occur very often in a given language. Words that are highly predictable from the context receive shorter fixations and more word skippings (see e.g., Balota, Pollatsek, & Rayner, 1985; Calvo & Meseguer, 2002; Kliegl et al., 2006; Rayner, 1998, 2009; Rayner, Ashby, Pollatsek, & Reichle, 2004). Many cognitive processes contribute to this effect, ranging from rather low-level priming effects to high-level language comprehension (see Rayner, 1998, for a review).

A second strategy to test assumptions on the interplay of different cognitive processes in reading has been to develop tasks, which involve similar visual and oculomotor processes as reading but differ with respect to the higher-level cognitive processing that is necessary to complete the task. In the *zzz*-string scanning task, originally introduced as *mindless reading* (Vitu, Oregan, Inhoff, & Topolski, 1995), participants read sentences in both their normal version as well as a transformed (or mindless) version where each letter is replaced with a *z* (see also Nuthmann, Engbert, & Kliegl, 2007; Rayner & Fischer, 1996). *z*-String scanning has similar visuo-oculomotor requirements as reading but shares none of the language-related processes. Mindless reading thus approximates reading without lexical and post-lexical processing (see Nuthmann & Engbert, 2009, for a simulation study).

In target-word search (Rayner & Fischer, 1996; Rayner & Raney, 1996), participants search through passages of text for a target word. All linguistic information, like word frequency and predictability of words, is present in the text. However, processing this information is not necessary to complete the task. Instead, the target can be detected based on superficial visual or orthographic analysis of words. Rayner and colleagues have investigated eye movements during target-word search and found no effect of word frequency on eye movements, contrary to robust frequency effects when reading the same text for comprehension. This finding suggests that lexical processing influences eyes movements during reading, but not in visual search for a target word.

Here, we combine these two approaches to add to our knowledge on eye-movement control in reading. We present a new paradigm, the reading of shuffled text, and we compare the influence of various variables on eye movements in this task to reading normal text. The basic idea underlying the shuffling of words is to convert meaningful sentences into meaningless word lists. We used the Potsdam Sentence Corpus (PSC), which consists of 144 single sentences (Kliegl, Grabner, Rolfs, & Engbert, 2004; Kliegl et al., 2006). Based on this sentence corpus, the order of words was randomly shuffled across the whole corpus, yielding randomly shuffled word lists, e.g.,

Affen Vorschlag Armen schmale Giebel Kanzler dem besser.
*Monkeys suggestion poor/arms narrow gable chancellor the better.*[1]

Jede ihrer Förster im Jahr Hunde meisten Gräfin Bauern.
*Each [of her/their] foresters [in the] year dogs most countess countrymen.*

In the randomization process, words were not shuffled within sentences, but for each word list words were randomly drawn from all original sentences in the PSC (cf., Morton, 1964, for a different approach to manipulate the context in English text). Readers were instructed to read these random lists of words. To ensure that readers would indeed process the shuffled and normal sentences, some trials were followed by a comprehension question or a word recognition probe. For shuffled word lists, participants were presented with a word triple and asked to indicate which word they recognized as part of the previous list; only content words were queried. For normal sentences, readers had to answer an easy three-alternative multiple-choice question pertaining to the content of the sentence.

How are eye movements controlled during reading of shuffled text? In the remainder of Section 1, we will derive specific predictions about how readers' eye movements might be affected by random shuffling of words. We will discuss: (1) basic visuomotor processes, (2) whether effects of lexical processing should occur, (3) differences in the predictability of words, (4) memory and post-lexical processes. Lastly (5) we will derive predictions about how theoretical models of reading can explain differences in word-frequency effects between normal and shuffled-text reading.

When reading shuffled text, low-level visuomotor requirements are similar to the ones in normal text reading. Therefore, similar visuomotor effects should be expected in eye movements. Linguistic information on single words, like their frequency, is also available in shuffled texts. Whether and to what degree this information will be relevant for eye guidance is unclear a priori and may depend on the strategy participants adopt to solve the task. In principle, superficial orthographic or phonological analysis can suffice to remember the words.[2] The use of such a strategy would predict that lexical processing does not influence eye movements in shuffled texts, similar to eye movements during target-word search (Rayner & Fischer, 1996; Rayner & Raney, 1996).

However, we expected that readers process words lexically when reading shuffled text and that this should affect their eye movements, in a similar manner as in normal reading. This is plausible (a) because lexical processing is highly automatic (see the Stroop effect, MacLeod, 1991) and (b) because readers were instructed to *read* the words (and not, for example, to *scan* them). In addition, (c) encoding the lexical identity of words should aid readers to do well in the word recognition queries and (d) readers may want to use post-lexical processing of, for example, semantic word information to memorize words. In sum, we expected that word frequency should affect eye-movement parameters during reading of shuffled text.

Further, we expected specific differences between the reading conditions. In randomly shuffled texts, upcoming words cannot be predicted based on their preceding context. Lacking word predictability should lead to a reduced word-skipping rate and increased fixation durations in reading of randomly shuffled texts compared to normal reading. This effect should be quite strong, because in normal text unpredictable words are often neighbored by predictable words, whereas in shuffled text none of the words are predictable. Although shuffled word lists are essentially free of meaning, readers may try to actively construct some meaning to better remember the words in the list (cf., Mason & Just, 2004; Myers, Shinjo, & Duffy, 1987). Also, we cannot exclude the possibility that some of the random word sequences may partially make sense and trigger automatic semantic or syntactic analyses. In the present study, however, we will focus on effects of lexical word processing, which is often assumed to be the primary cognitive process controlling eye movements during reading (e.g., Engbert et al., 2002, 2005; Reichle, Rayner, & Pollatsek, 2003; Reichle et al., 1998).

In any case, the shuffling of words does not only manipulate overall sentence meaning and the predictability of individual words, but is likely to affect other factors like the ease of retention of words. Shuffled text has no real meaning, which should make it more difficult to remember the words and may invoke different

[1] Note that languages differ from each other in various aspects. For example, nouns in German are always capitalized.

[2] Thanks to Keith Rayner for pointing this out.

memory-related processes than normal reading. These could contribute to a slower reading pace when reading shuffled text. Further, the specific instruction given to the participants, combined with the occasional word recognition probes, may cause differences in how readers construe their task when reading shuffled as opposed to normal text. Most importantly, only (low-frequency) content words are probed in the recognition test. It is possible that readers are aware of this and focus more strongly on the processing of salient low-frequency content words when reading shuffled text. In contrast, when reading normal text (high-frequency) function words and content words are equally important to construct meaning. We will outline more specific predictions that build upon this basic idea below.

To summarize, lexical processing should principally affect eye movements in both tasks. Therefore, we hypothesize that some basic mechanisms controlling the eyes when reading single unrelated words for recognition are not fundamentally different from the ones acting during normal text reading. However, post-lexical (especially memory-related) processes should differ between reading conditions. Task differences might lead to specific differences in how certain variables, most notably word frequency, modulate fixation times in shuffled text as opposed to normal reading. Such differences will be discussed on the basis of existing models of eye-movement control, with a focus on architectural principles embedded in our own SWIFT model (Engbert et al., 2005).

Cognitive models of eye guidance in reading make different assumptions about the nature of lexical processing and how attention is allocated to support such processing. According to sequential attention shift (SAS) models, most importantly the E-Z Reader model, attention is allocated serially to support lexical processing of only one word at a time (e.g., Reichle, Liversedge, Pollatsek, & Rayner, 2009; Reichle et al., 1998, 2003). Another group of models assumes guidance by a processing gradient (PG). In PG models, attention is distributed continuously as a gradient, which supports the processing of two or more words in parallel (e.g., Engbert et al., 2002, 2005; Reilly & Radach, 2006). Empirical support has been provided for both kinds of models, and aspects of the empirical findings and their theoretical implications are the subject of considerable debate (see e.g., Engbert & Kliegl, in press; Inhoff, Eiter, & Radach, 2005; Kliegl, Risse, & Laubrock, 2007; Kliegl et al., 2006; Pollatsek, Reichle, & Rayner, 2006a; Rayner, Pollatsek, Drieghe, Slattery, & Reichle, 2007; Rayner, White, Kambe, Miller, & Liversedge, 2003; Reichle, Liversedge et al., 2009).

As stated above, it could be that readers of shuffled text focus more strongly on the processing of low-frequency content words to better remember these words when reading shuffled text. Thus, shuffled text might influence allocation of attention during reading: It could change how the attentional gradient is dynamically modulated in response to foveal load (Henderson & Ferreira, 1990). In the following, we will outline this hypothesis in more detail. The *perceptual span* can be defined as the "region of the visual field from which useful information can be acquired during a given eye fixation" (Henderson & Ferreira, 1990, p. 417). It was studied in the *moving window paradigm* (McConkie & Rayner, 1975), where text is covered with a mask (e.g., XXX) and only the fixated words or letters are visible to the reader. The window of visible text moves with the readers' eyes, and covering parts of the text slows reading down. At a certain window size (about 14–15 letters to the right and 3–4 letters to the left), however, reading with a window proceeds at the same speed compared to when all text is visible, indicating the size of the perceptual span. In the SWIFT model (Engbert et al., 2005), the concept of a processing or attentional gradient combines the concept of a perceptual span with the notion of parallel processing of words in a sentence. The rationale here is that words within the perceptual span are processed in parallel, at rates decreasing with distance from the current fixation location.

*Does shuffling of words change the dynamical modulation of the perceptual span by foveal load?* The *foveal load hypothesis* (Henderson & Ferreira, 1990) postulates that the width of the perceptual span is modulated by foveal load (i.e., foveal processing difficulties). If foveal load is low the perceptual span is wide and attentional resources can be distributed across neighboring words. When foveal load increases, the perceptual span gets narrower and the resources left for processing parafoveal information decrease. Empirically, an incorrect preview for word $n + 1$ during fixations on word $n$ interferes with reading word $n + 1$ more strongly if word $n$ is of high-frequency, due to increased parafoveal processing in this condition (Henderson & Ferreira, 1990; see also Balota et al., 1985; Inhoff, Pollatsek, Posner, & Rayner, 1989; Inhoff & Rayner, 1986; Rayner & Pollatsek, 1987; White, Rayner, & Liversedge, 2005). In corpus analyses, the same mechanism is visible. Here, high-frequency words $n - 1$ increase preview for word $n$ during fixations on word $n - 1$. Because part of the processing of word $n$ could already be finished while still fixating word $n - 1$, the fixation on word $n$ is then shorter and the effect of frequency of word $n$ on fixation durations is weaker (Kliegl et al., 2006).

As outlined above, concerning its theoretical interpretation the foveal load hypothesis naturally adheres to the parallel processing assumption in reading. In PG models, low foveal load would lead to a widening of the attentional gradient. High foveal load, to the contrary, would narrow the attentional gradient such that only the fixated word would be processed. The basic foveal load finding (reduced preview benefit in case of increased foveal load) can also be accounted for within the SAS framework. The E-Z Reader model explains the effect by assuming that the second stage of lexical processing (L2) is a function of word frequency (Pollatsek, Reichle, & Rayner, 2006b; Reichle et al., 1998, 2003, 2006; Reingold & Rayner, 2006). In the model, L2 takes longer to complete for low-frequency words, which leads to less preview of the next word (and can even produce spill-over effects). Thus, the key signature finding of the foveal load hypothesis is compatible with both parallel and serial accounts of attention allocation during reading.

We now derive further, more specific predictions based on the assumption that foveal load modulates the perceptual span (Henderson & Ferreira, 1990) based on the PG framework. The basic assumption is that foveal load modulates the width of the attentional gradient. In addition, we assume that the processing resources are limited (i.e., that the total processing rate is constant at any time), such that capture of attentional resources by the fixated word would result in reduced processing of the neighboring word $n + 1$ (see Fig. 1 for an illustration).

As a first prediction for shuffled text, the effect of current-word frequency should be reduced if the modulation of the perceptual span by foveal load is strong, because low-frequency words capture more attentional resources compared to high-frequency words due to the contraction of the perceptual span. Second, a parallel processing account predicts that the influence of the upcoming word $n + 1$ on fixation durations depends on the frequency of the currently fixated word $n$. Because the amount of preprocessing of the next word depends on the width of the perceptual span (which in turn depends on the frequency of the fixated word), we expect parafovea-on-fovea effects to be modulated by foveal load (cf., Kliegl et al., 2006). Third, the current-word frequency effect should depend on the length of the currently fixated word $n$. A long word, be it of high or low-frequency, will fill more or less the whole perceptual span (Fig. 1a and b). Therefore, the current-word frequency effect should be fully visible. The effect might be weaker for short words, as they can benefit strongly from focusing of the perceptual span (Fig. 1c and d). Fourth, to the degree that short words $n$ benefit from focusing of the perceptual span, processing of successor words $n + 1$ should suffer from it. A short word $n$ with a low-frequency should attract all attentional resources. Accordingly, parafoveal processing of word

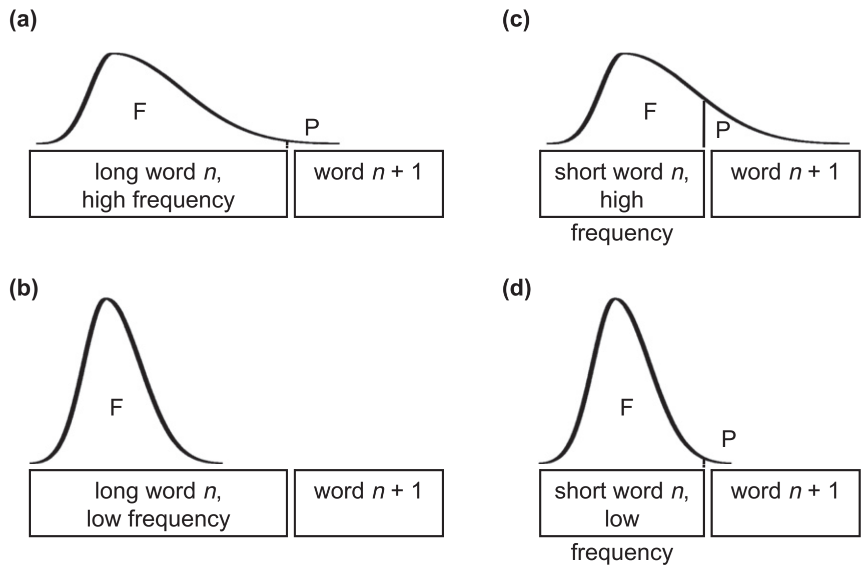


**Fig. 1.** Processing rate over foveal eccentricity; peak indicates fixation location. Predictions of the foveal load hypothesis for long words (left plots) vs. short words (right plots) with high (top row) vs. low (bottom row) frequency. Low word frequency equates to high foveal load. (1) Long word n: narrowing the perceptual span in response to a low-frequency word does not increase the processing resources available for the fixated word *n* (*F*) much (compare (b) with (a)). (2) Short word *n*: narrowing the perceptual span in response to a low-frequency word strongly increases the processing resources available for the fixated word *n* (compare (d) with (c)). *F* = processing resources available for the foveal word *n*; *P* = processing resources available for the parafoveal word *n* + 1.

*n* + 1 should be strongly reduced and fixation durations on word *n* + 1 should be enhanced (compared to good preview during a short high-frequency word *n*). A long word *n*, again, will fill the whole perceptual span independent of its frequency. For that reason, preprocessing of and fixation durations on word *n* + 1 should not strongly depend on the frequency of word *n*.

Deriving these four specific predictions is rather straightforward from the perspective of PG models supporting parallel word processing in reading. Notably, the predictions are derived based on one single mechanism, that is, the modulation of the perceptual span by foveal load.

## 2. Methods

### 2.1. Participants

Sixty university students participated in the study. Thirty readers took part in the shuffled reading condition. Their eye-movement data were compared with data generated by participants who read the Potsdam Sentence Corpus (PSC, normal sentence reading, *N* = 30), an age-matched subsample from a large set of data that has previously been reported in Kliegl et al. (2006). Both groups were tested in the same lab, using the same technical equipment. The two groups did not differ in age (shuffled-text reading: *M* = 22.8, *SD* = 3.4; normal reading: *M* = 22.6, *SD* = 3.6) and in psychometric tests of vocabulary (shuffled-text reading: *M* = 31.8, *SD* = 2.7; normal reading: *M* = 32.7, *SD* = 1.6), and digit-symbol substitution (shuffled-text reading: *M* = 61.7, *SD* = 9.6; normal reading: *M* = 59.2, *SD* = 9.4).

### 2.2. The Potsdam Sentence Corpus (PSC) and shuffled texts

The PSC comprises 144 German single sentences. They range from 5 to 11 words (*M* = 7.9, *SD* = 1.4), and there are 1138 words in total. Norms on psycholinguistic variables such as word length, printed word frequency (Geyken, 2006), and predictability norms from an independent cloze-task study are available for each word in the PSC. For details of materials and experimental procedure for the normal PSC data we refer to Kliegl et al. (2004, 2006).

To create shuffled text, each single sentence in the PSC was replaced by a shuffled word list. For each sentence, each word was replaced by a different word that was randomly drawn without replacement from the pool of all words that occur in the PSC. In this randomization procedure, the first word of an original PSC sentence was always the first word in a shuffled sentence; the same was true for the last words in sentences. All other words were drawn from random locations in a sentence. Using this constrained randomization procedure a separate set of 144 word lists was generated for each participant. As a consequence of this procedure, words in one word list were randomly drawn from many different sentences in the PSC.

### 2.3. Apparatus, materials and procedure

One group of participants read the original 144 PSC sentences, while the other group read a set of 144 random word lists. Sentences and word lists were presented in random order at a distance of 60 cm on the centerline of a 21-in. EYE-Q 650 Monitor (832 × 632 resolution; frame rate 75 Hz; font: regular New Courier 12; visual angle: 0.38° per character). A chinrest was used to minimize participants' head movements. Both eyes were monitored with an EyeLink II system (SR Research, Osgoode, ON, Canada) with a sampling rate of 500 Hz and an instrumental spatial resolution of 0.01°. Minimal head movements were corrected automatically by the EyeLink II system.

In order to motivate participants to read the word lists and/or sentences, simple questions occurred after 27% of the sentences and after one third of the word lists. In sentence reading, participants were asked questions pertaining to the meaning of the sentence. As response alternatives, a word triple was presented with the question and participants were required to indicate the correct word, which was always part of the sentence. In shuffled-text reading, participants were again presented with a word triple and were asked to decide which of the three words had been part of the list seen before. In both conditions, only nouns, verbs, or adjectives were queried in order to avoid changing the experiment into a (difficult) memory task. (Preliminary tests had shown that asking for prepositions, adverbs, etc. was difficult.) Participants were not informed about this particularity.

**Table 1**
Number of fixations for various types of fixations in shuffled and normal text reading.

| | | Shuffled text | Normal text | Total |
|---|---|---|---|---|
| 1 *N* of fixations | | 41,873 | 31,985 | 73,858 |
| 2 First/last word; first/last fixation | *N* | 11,075 | 9869 | 20,944 |
| | % | 26 | 31 | 28 |
| 3 Long fixation or amplitude | *N* | 195 | 118 | 313 |
| | % | 0.5 | 0.4 | 0.4 |
| 4 *N* of valid fixations | | 30,603 | 21,998 | 52,601 |
| 5 Not in first pass | *N* | 4575 | 2476 | 7051 |
| | % | 15 | 11 | 13 |
| 6 Different words | *N* | 2784 | 2931 | 5715 |
| | % | 9 | 13 | 11 |
| 7 Multiple fixations | *N* | 10,272 | 5130 | 15,402 |
| | % | 34 | 23 | 29 |
| 8 Single fixations | *N* | 12,972 | 11,461 | 24,433 |
| | % | 42 | 52 | 46 |

*Note*: Row 1 = 2 + 3 + 4; row 4 = 5 + 6 + 7 + 8. Data are from 30 readers in the shuffled, and 30 readers in the normal text condition. Data are from right eye.

### 2.4. Data selection

An initial screening excluded the records of sentences with blinks or loss of measurement from the data. Data from a maximum of 27 (*Median* = 3) sentences were excluded per participant. A binocular velocity-based algorithm for saccade detection (Engbert & Kliegl, 2003; Engbert & Mergenthaler, 2006) was used to identify saccades and fixations. To adjust for the reading situation, only fixations with a minimal duration of 10 ms and saccades with a minimal amplitude of 0.75° were detected. Fixations were assigned to letters within words. Sentences with less than three fixations and fixations left or right of the sentence borders were removed. This procedure resulted in a total number of 73,858 fixations (see Table 1 for separate numbers for the shuffled vs. normal text reading groups).

We excluded fixations according to the following criteria: (1) the first or last fixation in a sentence as well as fixations on the first or last word ($N = 20{,}944$), (2) fixations longer than 750 ms and fixations bordered by a saccade amplitude of 25 letters or longer ($N = 313$). The remaining fixations are valid fixations ($N = 52{,}601$). Among these we identified fixations that were not in first-pass reading[3] ($N = 7051$). Given that we wanted to examine influences from neighboring words, we only considered fixations where the left and right eye fixated on the same word. We thus excluded cases where the left and right eye fixated on different words ($N = 5715$; see Nuthmann & Kliegl, 2009, for an investigation of disparity between eyes). All measures of fixation durations or fixation probabilities were determined using the right eye. We further distinguished cases in which a word was fixated exactly once (binocularly reliable single fixation cases; $N = 24{,}433$) from cases in which a word had been fixated more than once during first-pass reading (multiple fixation cases; $N = 15{,}402$). In sum, the single fixation durations analyses reported below consider first-pass fixations where both eyes fixated on the same word.

In valid sentences, readers made first-pass fixations on a total of 55,323 words. When reading shuffled text, more words were fixated in first pass ($N = 29{,}704$) than when reading normal text ($N = 25{,}619$). For fixated words, all first-pass fixations were summed up to obtain gaze durations. For a given subject, words on which at least one invalid fixation (first/last word; first/last fixation; long fixation or saccade amplitude) was identified (shuffled: $N = 8372$; normal: $N = 8156$), as well as gaze durations that were longer than 1000 ms (shuffled: $N = 46$; normal: $N = 11$) were excluded from analysis. This procedure resulted in a total of 38,738 gaze durations (shuffled text: $N = 21{,}286$; normal text: $N = 17{,}452$).

[3] First-pass reading comprises all fixations on a word that occur before the first regression has originated from this word or a word following later in the sentence.

## 3. Results

### 3.1. Global summary statistics

Reading shuffled text resulted in a higher overall number of fixations than reading normal text. This also translated into a higher number of valid as well as first-pass fixations (see Table 1), and also more valid gaze durations. Accordingly, readers of shuffled text made more fixations per trial than normal text readers [10.1 vs. 7.8; $t(51) = 5.14$, $p < 0.001$; see Appendix A, Table A1, for descriptive statistics of eye movements]. Amplitudes for forward saccades were on average shorter in shuffled-text reading as compared to normal reading [6.1 vs. 7.6 letters; $t(55) = -5.85$, $p < 0.001$; see Fig. A1b) for the corresponding distributions of saccade lengths]. Shorter saccade lengths in shuffled text compared to sentence reading were associated with a strong reduction of skipping rate [0.10 vs. 0.21; $t(55) = -6.45$, $p < 0.001$] and an increase in refixation probability [0.16 vs. 0.08; $t(48) = 5.41$, $p < 0.001$]. Refixations were not only more frequent in the shuffled text condition but they were also more often rightward-oriented [90% vs. 79% of refixations in first-pass reading; $t(58) = 3.2$; $p < .01$]. The decrease in skipping probability and increase in refixation probability canceled each other out such that the probability of single fixation was similar for the two groups [0.70 vs. 0.67; $t(58) = 1.69$; $p = 0.10$].

The percentage of regressions was exactly the same (0.06 vs. 0.06). Likewise, the distribution of backward-oriented saccade amplitudes did not differ between reading conditions [Fig. A1b]. The number of fixations in second- and more-pass reading was largely enhanced in shuffled-text reading (4575 vs. 2476 fixations; $\chi 2(1) = 625$; $p < .001$; see Table 1).

In shuffled-text reading, readers initially fixated further to the left in a word compared to normal text reading. This difference was significant for single fixation cases [initial fixation on letter 2.5 vs. 2.7; $t(53) = -3.22$, $p < 0.01$], while there was a trend for the first of multiple fixations [letter 2.0 vs. 2.2; $t(53) = -1.85$, $p = 0.07$].

Fixation durations were generally longer in readers of shuffled compared to normal text. This effect showed as a moderate shift in mean and skew in the corresponding global fixation duration distribution [Fig. A1a]. The difference in fixation durations was observed across all types of fixations; it was significant for single [254 vs. 213 ms; $t(55) = 4.37$, $p < 0.001$], first [227 vs. 199 ms; $t(58) = 3.50$, $p < 0.001$], and second [197 vs. 172 ms; $t(58) = 2.74$, $p < 0.01$] fixations, as well as for gaze durations [293 vs. 231 ms; $t(50) = 4.89$, $p < 0.001$]. As a result of the higher number of fixations and the longer fixation durations, the reading rate was strongly reduced in readers of shuffled text as compared to normal text.

Memory performance was close to perfect for readers of normal text (97.5% of the questions, $SD = 3.6$, were answered correctly). Readers of the shuffled text answered 85% of the questions correctly ($SD = 3.1$).

### 3.2. Linear mixed-effects models

We used gaze duration and single fixation duration as dependent measures in our analyses. Gaze durations and single fixation durations were log-transformed to avoid problems with heteroscedasticity. To determine the impact of various predictors on log-fixation durations in shuffled text vs. sentence reading, a linear mixed-effects model (LME; e.g., Baayen, 2008; Baayen, Davidson, & Bates, 2008; Pinheiro & Bates, 2000; see also Kliegl, 2007; Kliegl, Masson, & Richter, 2010; Kliegl et al., 2007) was tested, using the *lmer* program of the *lme4* package (Bates & Sakar, 2008). Plots were created using the *ggplot2* package (Wickham, 2009). The packages and programs are supplied in the *R* system for statistical computing (R Development Core Team, 2008; under the GNU General Public License, Version 2, June 1991).

Fixed effects in LME terminology correspond to regression coefficients in standard linear regression models. They can also estimate slopes or differences between conditions. A number of fixed effects were entered into the model. We tested the influence of visual and lexical factors characterizing the currently fixated word $n$ by including its length (i.e., 1/length) and its frequency, with linear and quadratic (cf., Kliegl, 2007) effects, as well as their multiplicative interaction (cf., Pollatsek, Reichle, Juhasz, Machacek, & Rayner, 2008; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999). To test for lag effects of the previous word $n-1$ on fixation durations on the fixated word $n$, we used word $n-1$ length (1/length; cf., Pollatsek et al., 2008) and frequency as predictors (cf., Rayner & Duffy, 1986; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999). Likewise, successor effects were tested by including word $n+1$ length (1/length) and frequency (cf., e.g., Kennedy & Pynte, 2005; Vitu, Brysbaert, & Lancelin, 2004). We further added the length of the incoming (cf., Pollatsek, Rayner, & Balota, 1986; Radach & Heller, 2000; Vitu, McConkie, Kerr, & O'Regan, 2001) and outgoing saccades as model predictors. To capture the inverted-optimal viewing position effect for fixation durations (IOVP, Nuthmann, Engbert, & Kliegl, 2005; Nuthmann et al., 2007; Vitu, Lancelin, & d'Unienville, 2007; Vitu, McConkie, Kerr, & O'Regan, 2001) the relative fixation position within a word (i.e., fixated letter number divided by word length) was included as a linear and as a quadratic effect.

In addition, three further predictors involving multiplicative interaction terms of continuous variables were added to the model. We tested whether the influence of current-word frequency was modulated by the frequency of the prior word (a prediction derived from the foveal load hypothesis, Henderson & Ferreira, 1990; Kliegl et al., 2006). Likewise, we examined whether the influence of the frequency of the parafoveal word $n+1$ depended on limits of visual acuity (i.e., on the length of the fixated word; cf., Kennedy & Pynte, 2005) and on attentional constraints (i.e., on the frequency of the fixated word; a second prediction derived from the foveal load hypothesis). Except for the quadratic effect of current-word frequency (Kliegl, 2007) and lacking effects of word predictability, this set of predictors was identical to the set of predictors tested with repeated measures multiple regression analysis (*rmMRA*) reported by Kliegl et al. (2006; see also several random-subject *lme* models in Kliegl (2007)).

For statistical modeling we used two complementary approaches. First, we tested whether the fixation-level fixed effects differed between the shuffled and the normal reading group (i.e., we tested cross-level interactions). This was done by simultaneously including all of the fixation-level effects as well as their interactions. *Experimental condition* was included as a dummy factor, using the shuffled text condition as the reference group.[4] In addition, we estimated how strongly mean fixation durations varied with participants and words by fitting crossed random intercepts for participants and words (if the same word occurred more than once in the corpus, the same random effect was used for all of these occurrences, yielding unique word ID). Instead of estimating a slope or a difference between conditions, random effects estimate the variance that is associated with the levels of a certain factor. After including these effects into the model, non-significant predictors were dropped. The results for this final model are reported in the text below; for an overview see Appendix B, Table B1. Values of $t > 1.96$ indicate significance of a predictor, while effects with $t > 1.645$ indicate marginal significance. Second, we tested whether the fixation-level fixed effects described above are significant in each of the reading conditions separately. To do so, we included each of these predictors twice within one model: once nested under shuffled and once nested under normal text reading.[5] In this post hoc model, we again used the same random effects and the same procedure for dropping predictors. In the following we report the effects of word frequency when reading normal and shuffled text.

#### 3.2.1. Effects of current-word frequency

*3.2.1.1. Main effect of word frequency.* The word-frequency effect on fixation durations is one of the most basic and best-replicated findings in reading research: low-frequency words are fixated longer than high-frequency words (e.g., Inhoff & Rayner, 1986). Accordingly, fixation durations should decrease with increasing current-word frequency. Such an inverse relationship will be referred to as a *negative effect* of a variable (indicated by a negative fixed effect coefficient), while we will use the term *positive effect* (with a positive coefficient in the model) for cases in which fixation durations increase with higher values in the predictor variable.

For the log-frequency of the fixated word $n$, we found the expected negative influence on gaze durations (see Fig. 2; for normal sentence reading: $b = -0.032$, $SE = 0.006$, $t = -5.23$). For readers of shuffled text, however, the linear effect of word frequency disappeared (in Fig. 2, low-frequency words show somewhat longer gaze durations because word frequency is confounded with effects of word length. The LME model controls for such effects and reveals a null-effect of word frequency: ($b = 0.004$, $SE = 0.066$, $t = 0.67$; for the difference between conditions: $b = -0.026$, $SE = 0.004$, $t = -6.7$). The quadratic current-word frequency effect did not significantly differ between the two conditions and was overall significant ($b = 0.020$, $SE = 0.005$, $t = 4.3$).

For single fixation durations, the linear current-word frequency effect also significantly differed between the two reading conditions ($b = -0.028$, $SE = 0.004$, $t = -7.0$). Like in gaze durations, it

[4] Consequently, if the interaction of a fixation-level fixed effect with experimental condition is kept in the model (e.g., frequency of word $n *$ experimental condition), the coefficient estimating the fixation-level fixed effect itself (i.e., in this case the main effect of frequency of word $n$) tests the influence of this variable in the shuffled text condition. If the same interaction is, however, removed from the model because it does not reach significance, the fixation-level fixed effect (e.g., the main effect frq. $n$) represents the average effect of the variable (frq. $n$) for both reading conditions.

[5] Nesting a covariate (e.g., word frequency) under the level of an experimental factor (e.g., under shuffled-text reading) can be done by means of setting all values of the covariate for the other factor levels (in this case for normal sentence reading) to zero and to center the covariate within the critical factor level. As a result, the effect of the covariate is estimated and tested only within the specified factor level (i.e., the frequency effect among readers of shuffled text; cf., Kliegl, 2007).

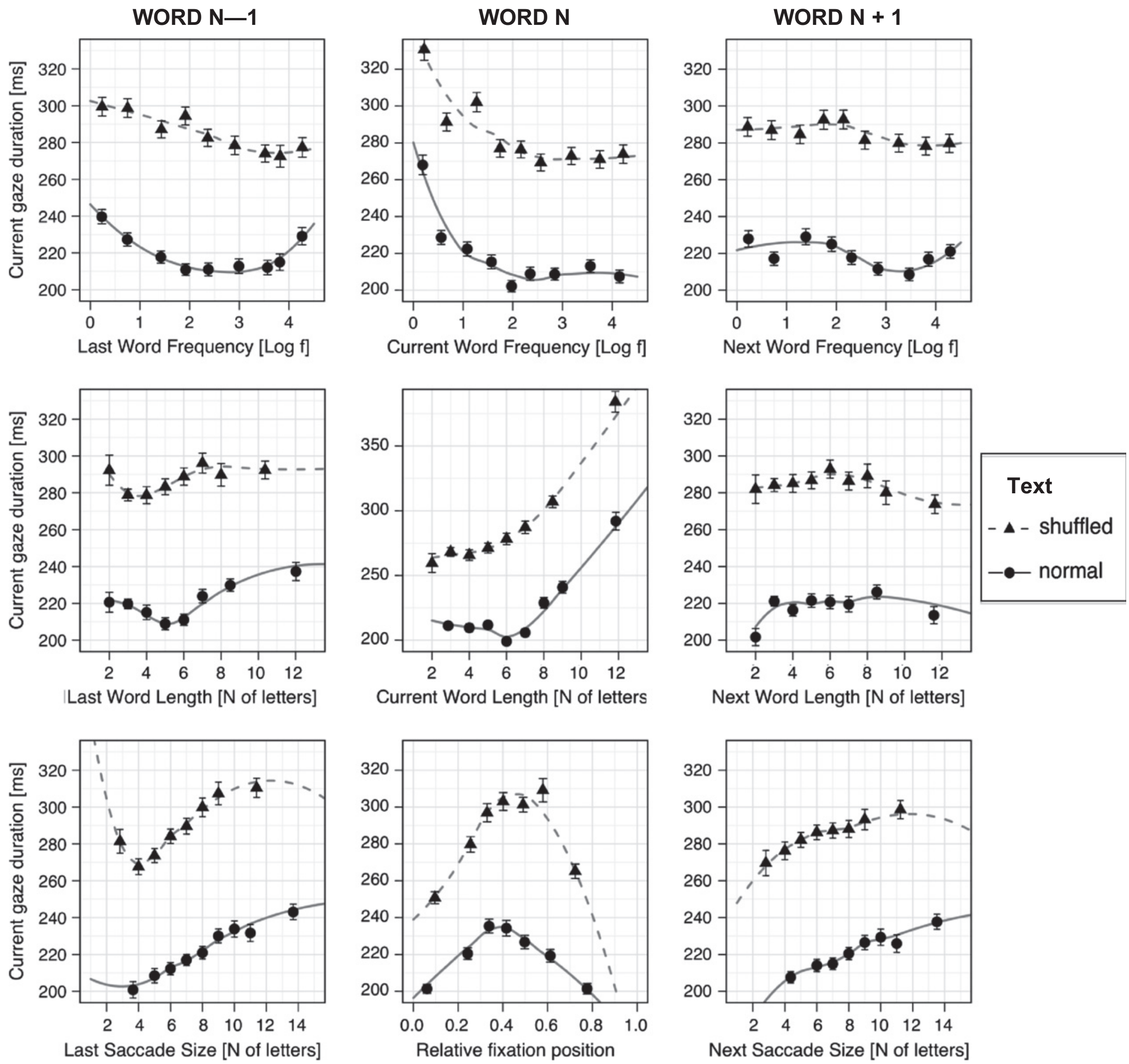


**Fig. 2.** Nine main effects for *gaze durations* on word *n* for reading shuffled (*N* = 30, triangles and dashed lines) and normal text (*N* = 30, circles and solid lines). Predictors are frequency and length of words $n-1$, $n$, and $n+1$ (first two rows), the amplitude of the incoming saccade, the relative fixation position (rfp) in the word (linear + quadratic trend), and the amplitude of the outgoing saccade (last row). For each predictor, fixations were binned into categories with a minimum of 800 fixations. Error bars are within-subject 95% confidence intervals (using the method described by Cousineau, 2005). In addition, the predictions from a least squares local regression model, applied to the full set of ungrouped data, are plotted for each effect.

was weaker in readers of shuffled text. However, the effect actually changed its sign for single fixation durations. In sentence readers, low-frequency words were fixated significantly longer than high-frequency words ($b = -0.028$, $SE = 0.005$, $t = -5.8$). For readers of shuffled text, this traditional negative frequency effect numerically turned positive, such that low-frequency words were fixated for less time than high-frequency words. This positive frequency effect was marginally significant ($b = 0.010$, $SE = 0.005$, $t = 1.84$). The quadratic frequency effect on single fixation durations did not significantly differ between the two conditions and was overall significant ($b = 0.017$, $SE = 0.004$, $t = 4.5$).

*3.2.1.2. Interaction of word frequency and word length.* The current-word frequency effect on log gaze durations was modulated by word length, as there was a stronger frequency effect for long compared to short words. This was true for normal reading (see also Kliegl et al., 2006) as well as reading of shuffled text (for the overall interaction of word length and frequency: $b = 0.401$, $SE = 0.055$, $t = 7.3$). This interaction did not significantly differ between the two reading groups. For sentence readers, the current-word frequency effect was negative for both long and short words (Fig. 3a). For readers of shuffled text, this effect changed its sign. Low-frequency words were actually fixated shorter than high-frequency words, if the words were of short length. These word length dependent linear frequency effects were significant in a post hoc analysis.[6]

[6] To test these effects, the word length-variable was dichotomized (median-split; short words had five letters or less) and word frequency was nested under long and under short words in the shuffled and in the normal text reading condition (yielding four linear effects of word frequency for these four conditions). The new current-word length and frequency variables were used in an additional post-hoc mixed-effects model that lacked the overall linear effects of word length, frequency, and their interaction, and that was otherwise identical to the first post-hoc model (i.e., testing fixation-level effects nested under experimental condition). The linear frequency effect was significantly negative in three conditions [in short ($b = -0.013$; $SE = 0.007$; $t = -1.98$) and long ($b = -0.071$; $SE = 0.010$; $t = -7.2$) words for normal text reading and in long words ($b = -0.040$; $SE = 0.011$; $t = -3.8$) for shuffled-text reading], but was significantly positive for short words among readers of shuffled text ($b = 0.013$, $SE = 0.006$, $t = 2.0$) [effect of word length (short vs. long words): $b = 0.098$; $SE = 0.011$; $t = 8.9$; (Exp) $*$ (word length): $b = -0.046$; $SE = 0.009$; $t = -5.3$].

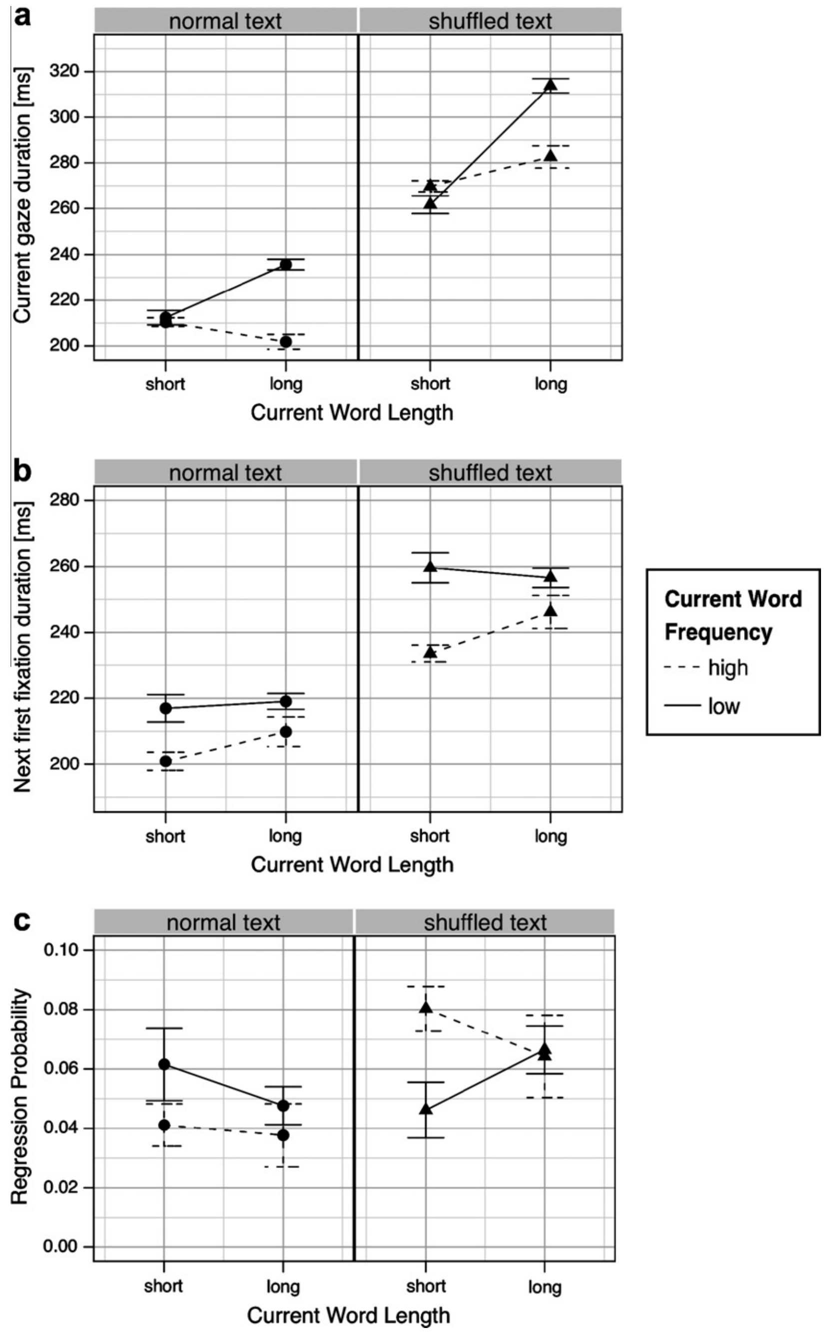


**Fig. 3.** Interaction between length and frequency of word *n* for normal (left plots, circles) vs. shuffled (right plots, triangles) text reading. (a) Effects on *gaze duration* on word *n*. (b) Modulation of *first fixation duration* on word *n* + 1, defined as the duration of the next fixation after having made one first-pass single fixation (see Section 2 for selection criteria) on word *n* and given that this next fixation is on word *n* + 1. (c) Effects on *regression probability* to word *n*, defined as the probability of regressing to word *n* after having made one first-pass single fixation on word *n* and one fixation on word *n* + 1. Short words are five or fewer letters long; DWDS frequencies were split on medians (calculated across both groups). Error bars are within-subject 95% confidence intervals (using the method described by Cousineau, 2005).

While the current-word frequency effect on log single fixation durations was significantly modulated by word length for readers of normal sentences ($b = 0.160$, $SE = 0.041$, $t = 3.9$), this modulation was not significant for participants reading shuffled text ($t = -1.38$; for the condition-difference: $b = 0.177$, $SE = 0.037$, $t = 4.8$). However, we again tested the same post hoc model as reported for the corresponding interaction in the gaze duration analysis and again found current-word frequency effects to be significantly positive only for short words among readers of shuffled text ($b = 0.015$, $SE = 0.005$, $t = 2.95$). In normal sentence reading, however, the frequency effects were significantly negative in both word length conditions ($b$s < −0.011, $t$s < −2.0).

To summarize, during reading of shuffled text the current-word frequency effect on gaze and single fixation durations was overall strongly reduced. It disappeared for gaze durations and was actually reversed for gaze durations on short words and for single fixation durations, yielding longer fixations on high- compared to low-frequency words. However, the standard effect of word frequency, with longer fixations on low-frequency words, was observed on long words in the gaze duration analysis. Also, the quadratic effect of word frequency was present during reading of normal as well as shuffled text.

#### 3.2.2. Effects of distributed processing: lag and successor frequency

*3.2.2.1. Lag effects.* The effect of the frequency of word $n-1$ on gaze durations did not significantly differ between the two groups of readers ($t=-1.2$). It was significant and negative in both groups (shuffled text: $b=-0.033$, $SE=0.003$, $t=-13.2$; normal text: $b=-0.040$, $SE=0.003$, $t=-11.9$). For single fixation durations, the lag effect of word $n-1$ frequency was numerically weaker in readers of shuffled text. However, the condition-difference for the slope of word $n-1$ frequency only approached significance ($t=-0.008$, $SE=0.004$, $t=-1.7$). The effect was still strong and highly reliable in readers of shuffled text ($t=-0.034$, $SE=0.003$, $t=-12.5$).

*3.2.2.2. Successor effects.* The effect of the frequency of the upcoming word $n+1$ on gaze durations did not significantly differ between shuffled and normal text reading, however there was a trend towards a stronger effect in shuffled text readers ($b=0.006$, $SE=0.003$, $t=1.86$). Gaze durations were generally shorter before high-frequent words $n+1$ (shuffled PSC: $b=-0.015$, $SE=0.003$, $t=-5.9$; normal PSC: $b=-0.011$, $SE=0.003$, $t=-3.3$). The same was true for the successor effect on single fixation durations: there was a significant effect for shuffled ($b=-0.015$, $SE=0.003$, $t=-5.5$) and for normal text readers ($b=-0.010$, $SE=0.003$, $t=-3.2$), but no significant slope-difference ($b=0.005$, $SE=0.003$, $t=1.6$). Thus, we found strong, consistent, and highly reliable effects of lag and successor-word frequency on gaze and single fixation durations during normal and shuffled-text reading.

#### 3.2.3. Interactions of frequencies of neighboring words

*3.2.3.1. Lag effects.* For gaze durations, the interaction between word $n$ and word $n-1$ frequency was significant in the normal sentence reading condition ($b=-0.004$, $SE=0.002$, $t=-2.1$). It was also significant for readers of shuffled text ($b=0.015$, $SE=0.001$, $t=10.4$). However, the coefficient was opposite in sign and higher in absolute value (for the difference: $b=-0.019$, $SE=0.002$, $t=-7.9$). Among readers of shuffled text, gaze durations were especially prolonged if word $n$ and word $n-1$ were both low in frequency (see Fig. 4a). For readers of normal text, on the other hand, gaze durations were particularly shortened in the case of high-frequent words $n$ and $n-1$.

For single fixation durations, we also found a strong and highly significant interaction between word $n$ and word $n-1$ frequency for readers of the shuffled PSC (i.e., a foveal load lag effect: $b=0.019$, $SE=0.002$, $t=12.3$). This interaction was significantly stronger ($b=-0.018$, $SE=0.002$, $t=-7.2$) than the corresponding interaction for normal text.[7] As for gaze durations, the lag-frequency effect was stronger in low- than in high-frequency words $n$ (see Fig. 5a).

*3.2.3.2. Successor effects.* The interaction between word $n$ and word $n+1$ frequency on gaze durations was not significant for readers of normal sentences, replicating prior research (Henderson & Ferreira, 1993; Kennison & Clifton, 1995; Kliegl et al., 2006; White et al., 2005). However, we observed a significant interaction in the shuffled-text reading condition ($b=0.010$, $SE=0.002$, $t=5.4$): The frequency effect of word $n+1$ on gaze durations was stronger if word $n$ was a low-frequency word (see Fig. 4b).

Similarly to the gaze duration data, the interaction of frequency of word $n$ and word $n+1$ was significant in the shuffled ($b=0.010$, $SE=0.002$, $t=4.9$) but not in the normal text reading condition ($t=-0.55$; condition-difference: $b=-0.010$, $SE=0.002$, $t=-4.0$) when analyzing single fixation durations. For the shuffled text readers, the parafovea-on-fovea effect of word $n+1$ frequency on single fixation durations was negative (i.e., longer fixation durations next to low-frequent words $n+1$) if the foveal word had a low-frequency. Surprisingly this effect numerically turned positive for high-frequent words $n$ (i.e., shorter fixation durations next to low-frequent words $n+1$; see Fig. 5b).

In summary, foveal load effects were much stronger in readers of shuffled text. In particular, the frequency of the last word $n-1$ modulated effects of current-word frequency more strongly, and the current-word frequency modulated successor-frequency effects when reading shuffled text.

### 3.3. Further tests of relative word-frequency effects

#### 3.3.1. Relative Lag-frequency effects – fixation durations

If the preview of word $n+1$ during fixations on word $n$ depends on the interaction of word $n$ length and frequency, then increased preview should show in shorter fixations on the next word $n+1$ (i.e., in a reduced spill-over effect). To test this, we refit the primary linear mixed model described above to regress the (log) duration of the first fixation on word $n+1$ after having made a single fixation on word $n$ on all the predictors reported above. In addition, we added the lag-frequency times lag-word length interaction to the set of fixed effects (note that these lag effects correspond to the current-word frequency and length effects in the previous models). Cases in which word $n+1$ was skipped during first-pass reading were excluded from the analysis, resulting in a total of 16,577 fixations. While there was a highly significant interaction of word $n$ length $\times$ frequency on fixation durations on word $n+1$ for readers of shuffled text ($b=-0.179$, $SE=0.033$, $t=-5.4$), this interaction was significantly weaker for readers of normal text ($b=0.119$, $SE=0.055$, $t=2.2$). As can be seen in Fig. 3b, high-frequency words $n$ lead to shorter fixation durations on word $n+1$ and this frequency-based preview benefit effect was significantly stronger for short compared to long words $n$. This was particularly the case for readers of shuffled text.

#### 3.3.2. Relative Lag-frequency effects – regression probability

To follow up on the reversed frequency effects for short words, we tested how word $n$ length, word $n$ frequency, and their interaction influenced the probability of regressing back to word $n$ after having fixated word $n+1$ once. We fitted a generalized (logistic) linear mixed model using regressions from word $n+1$ to word $n$ (after a first-pass single fixation on word $n$ and one fixation on word $n+1$) as the binary dependent variable ($N=21{,}129$ fixations). Predictors in the model were word $n$ frequency and length (i.e., $1/wl$), their interaction, frequency of word $n+1$, as well as interactions of these variables with experimental condition (shuffled vs. normal PSC readers) using crossed random intercepts over subjects and over unique word id.

The effect of word $n$ frequency (i.e., of the regression target) on regression probability significantly differed between shuffled and normal PSC reading ($b=-0.35$, $SE=0.07$, $p<.001$). Readers of

[7] The interaction of word $n$ and word $n-1$ frequency was not significant for the normal PSC reading sample that we used in this study ($t=0.49$). However, this same interaction has earlier been found to be highly reliable across various samples of participants reading the PSC (see Kliegl, 2007; Kliegl et al., 2006). Therefore, we checked whether the interaction that we found for shuffled text readers was also stronger than the corresponding effect in other samples reading the normal PSC. To do so, we fitted a linear mixed-effects model using the same predictors as the ones reported in Kliegl (2007, Table 1) using non-transformed single fixation durations. We then checked whether the interaction-coefficient for shuffled PSC readers was larger than equivalent coefficients for other samples reading the normal PSC. The comparison with the data reported in Kliegl (2007, Table 1) reveals that the largest coefficient for this interaction in any of the other PSC samples was $b=3.0$ and was thus more than two standard errors below the coefficient that we found for the shuffled reading group ($b=4.5$, $SE=0.41$, $t=11.2$). Thus, the interaction of word $n$ and word $n-1$ frequency was stronger in readers of shuffled text compared to many observed samples of participants reading the normal PSC.

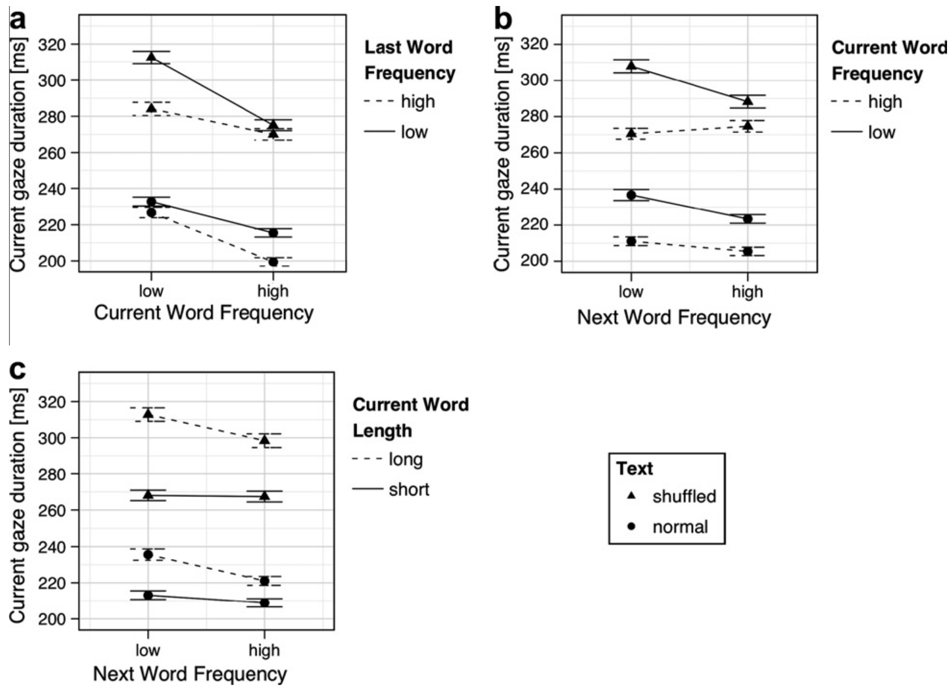


**Fig. 4.** Modulation of *gaze durations* on word *n* due to three interactions for readers of normal (circles) and shuffled (triangles) text: (a) frequency of word $n *$ frequency of word $n - 1$, (b) frequency of word $n + 1 *$ frequency of word *n* and (c) frequency of word $n + 1 *$ length of word *n*. Dependent variable is always gaze duration on word *n*. Short words are five or fewer letters long; DWDS frequency were split on medians. Error bars are within-subject 95% confidence intervals.

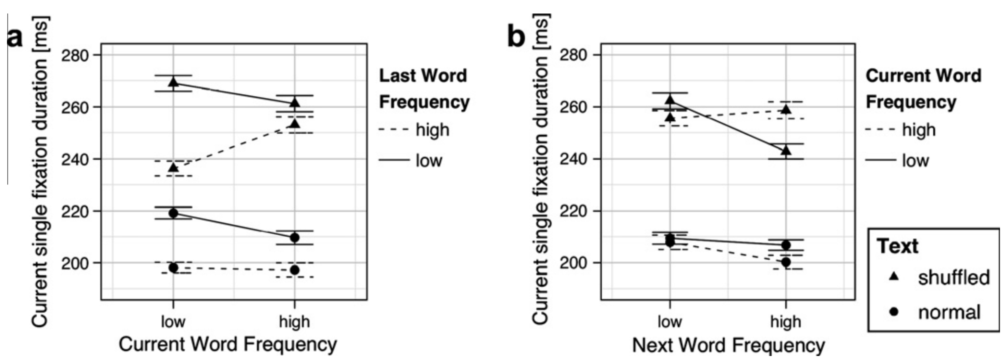


**Fig. 5.** Modulation of *single fixation durations* on word *n* due to two interactions for readers of normal (circles) and shuffled (triangles) text: (a) frequency of word $n *$ frequency of word $n - 1$ and (b) frequency of word $n + 1 *$ frequency of word *n*. Dependent variable is always single fixation duration on word *n*. Short words are five or fewer letters long; DWDS frequency were split on medians (calculated across both groups). Error bars are within-subject 95% confidence intervals (using the method described by Cousineau, 2005).

normal text made significantly more regressions to low-frequency compared to high-frequency words (i.e., a negative frequency effect; $b = -0.17$, $SE = 0.05$, $p < .001$). Readers of shuffled text, to the contrary, made significantly more regressions to high-frequency compared to low-frequency words (i.e., a positive frequency effect; $b = 0.18$, $SE = 0.05$, $p < .001$; see Fig. 3c). We further tested how the word *n* frequency effect depended on word length in the two reading conditions and found a marginally significant interaction for readers of shuffled text ($b = 0.85$, $SE = 0.48$, $p = .08$). Post hoc tests revealed that readers of shuffled text made more regressions to short high-frequency compared to short low-frequency words (i.e., a positive frequency effect for short words; $b = 0.24$, $SE = 0.06$, $p < 0.001$) while this frequency effect was not significant for long words *n* ($p = .27$). In readers of the normal PSC the frequency effect did not depend on word length ($p = .26$). Note that differences in word $n + 1$ frequency between reading conditions cannot be the source of these effects because this was statistically controlled for in the regression model.

## 4. Discussion

Eye movements in reading are affected by both low-level visual and oculomotor factors as well as higher-level cognition related to language processing. With the present work we introduce the shuffled-text reading task as a new paradigm to investigate the interplay of low-level and high-level factors in reading. In the reported experiment, the words of a well-investigated corpus of single sentences (PSC, Kliegl et al., 2004, 2006) were randomly shuffled to create meaningless word lists. For each shuffled sentence, words from different original sentences were randomly selected. Participants' task was to read the presented text. To ensure that participants complied with the instructions, about a third of the trials were followed by a comprehension question (normal sentences) or a word recognition probe (shuffled word lists).

The eye movements of participants reading these shuffled meaningless sentences were compared with those from participants who read the normal meaningful PSC sentences. A detailed statistical analysis of variables known to modulate fixation times (cf., Kliegl, 2007; Kliegl et al., 2006) showed various similarities and differences between the two tasks. Overall, our predictions as outlined in the Introduction were supported by the experimental results.

First, there was a considerable degree of similarity in the eye movements between readers of shuffled and normal text. We investigated how seven visuomotor variables influenced single fixation durations: the length of the fixated word $n$, the length of the last word $n-1$, and the length of the next word $n+1$, the amplitudes of the incoming and outgoing saccades, and the slope and location of the fixation-duration inverted-optimal viewing position (IOVP) effect. We found no evidence that these influences on single fixation durations differed between readers of shuffled and normal text, with only one marginal difference for the length of word $n-1$. This finding is consistent with our assumption that similar visual and oculomotor processes were in place when reading shuffled and normal text.

Second, there was no current-word frequency main effect on fixation times when reading shuffled text. This is surprising, but in line with work by Rayner and colleagues who found no effect of word frequency on eye movements in a task where participants searched for a target word in normal text (Rayner & Fischer, 1996; Rayner & Raney, 1996). The absence of word-frequency effects in visual search suggests that lexical word processing does not influence eye movements in this task. Was lexical processing also irrelevant for eye guidance when reading shuffled text? Although low-frequency words did not receive longer fixations than high-frequency words overall, we nevertheless found several strong and expected effects of word frequency on fixation durations during shuffled-text reading. In particular, effects of distributed processing, i.e., the influence of lag- and successor-word frequency, the quadratic effect of current-word frequency (Kliegl, 2007), and the coefficient for the interaction of current-word frequency with word length were highly reliable and more or less unchanged during reading of shuffled as compared to reading of normal text. Overall, low-frequency words were not looked at longer when reading shuffled text. However, this standard effect of word frequency was present for long words (see Fig. 3). At the same time, we found reversed effects of current-word frequency on gaze durations for short words (Fig. 3) and on single fixation durations. In these cases, fixations were longer on high- than on low-frequency words, which is opposite to what is found in normal reading. Taken together, these effects suggest that readers of shuffled text processed words lexically and that lexical word processing influenced their eye movements.

Notably, the probability of making a between-word regression as well as the distributions of leftward-oriented saccades were virtually identical for the two reading conditions [see Appendix A, Fig. A1b]. This striking agreement in distributions is well in line with the notion that most regressive eye movements when reading easy normal sentences like the PSC are triggered by unfinished word recognition (cf., Engbert et al., 2005; Nuthmann & Engbert, 2009). However, we cannot exclude the possibility that additional post-lexical processes, assuming that they might occur in one way or the other when reading shuffled text, may trigger the same amount and the same distribution of regressive eye movements in both tasks.

Third, we found support for our predictions with regard to slower processing. All measures of fixation durations (single, first, and second fixations as well as gaze durations) were significantly increased when reading shuffled as compared to normal text. Also, we observed a reduced skipping rate along with a strong increase in refixation probability. First and foremost, we attribute these results to the fact that the shuffling procedure removes the predictability of words. In addition, post-lexical integration and memorization of words should be harder in shuffled text, potentially contributing to the slower reading speed. In particular, the observed increase in second- and more-pass reading fixations may reflect active attempts of readers to try and memorize words and/or understand meaningless shuffled text. Another effect hinting towards memorization processes in shuffled-text reading is the stronger effect of word length as compared to normal reading. As longer words take more time to encode phonologically (Baddeley, Thomson, & Buchanan, 1975), the stronger word length effect would be in line with the idea that readers encode words in the phonological loop when reading shuffled text.

Further in-depth analyses revealed very specific processing differences between the two tasks. We argue that the reported pattern of results supports the hypothesis that the perceptual span was more strongly modulated by foveal load among readers of shuffled text compared to readers of normal sentences. In the following, we provide a detailed discussion of the results with respect to distributed processing (Section 4.1), the modulation of the perceptual span (Section 4.2), alternative explanations for changed frequency effects (Section 4.3), and PG vs. SAS models (Section 4.4).

### *4.1. Replication of effects of distributed processing*

Recently, Kliegl et al. (2006) used corpus analyses to investigate the influence of the foveal word $n$ as well as of neighboring words $n-1$ and $n+1$ on fixation durations on word $n$. They reported strong and consistent parafovea-on-fovea effects, yet their validity has been questioned (Rayner et al., 2007; but see Kliegl, 2007). Much of the criticism pertained to the correlational nature of the reported lag and successor effects. Here, we counter this argument by reporting robust and highly reliable effects of distributed processing for readers of shuffled text. When creating the shuffled word lists, each word was selected at random from all words in the corpus, and this random selection was done for each participant separately. Thus, observed effects are experimental in nature and allow the conclusion that processing neighboring words $n-1$ and $n+1$ causally affected fixation durations on the fixated word $n$. The effects of neighboring words on fixation durations on word $n$ were highly similar in normal and shuffled-text reading. This (a) suggests that these effects generalize to other reading situations, and (b) supports the validity of these effects in normal sentence reading.

### *4.2. A stronger modulation of the perceptual span in shuffled-text reading*

Our prediction was that readers of shuffled text should primarily focus on the processing of salient low-frequency content words to better remember them for the recognition task. From a perspective of a theoretical framework supporting parallel processing of words in the perceptual span (e.g., SWIFT, Engbert et al., 2005), such a strategy predicts a stronger modulation of the perceptual span by foveal load (Henderson & Ferreira, 1990) during reading of shuffled text. This prediction was supported by our findings.

#### *4.2.1. Relative lag effect*

The primary prediction derived from the foveal load hypothesis (Henderson & Ferreira, 1990; see also Balota et al., 1985; Inhoff & Rayner, 1986; Inhoff et al., 1989; Rayner & Pollatsek, 1987) states that the difficulty of a word $n-1$ (e.g., its frequency) modulates the amount of preview that is available for the next word $n$ during fixation on word $n-1$. High-frequency words $n-1$ would allow strong preprocessing of word $n$ during the previous fixation. This preview can be measured by the benefit of having seen a correct compared to an incorrect preview during the previous fixation (Henderson & Ferreira, 1990). In corpus analyses, extensive parafoveal preprocessing of word $n$ during fixations on word $n-1$ should attenuate the current-word frequency effect on word $n$. Previous words $n-1$ of low-frequency should result in a strong current-word frequency effect, while high-frequency words $n-1$ should go along with weaker current-word frequency effects (cf., Kliegl et al., 2006). In the present data, this interaction was stronger for readers of shuffled compared to normal text. We conclude that the modulation of the perceptual span is stronger in readers of shuffled text than in readers of normal text. Readers of shuffled text widen their perceptual span more strongly when fixating a word of high-frequency and focus their attention more strongly when reading a low-frequency word. To follow up on this hypothesis, we derived several qualitative predictions from a parallel model of word processing during reading (assuming that the total amount of processing resources is limited).

#### 4.2.2. Current-word frequency effects

The data supported the prediction that current-word frequency effects should be weaker if the modulation is stronger. In fact, when reading shuffled text, the frequency effect completely disappeared (gaze durations) or even turned into a small positive effect (single fixation durations). This is a noteworthy finding, because the negative word-frequency effect for fixation times (longer fixations on low-frequency than on high-frequency words) is one of the cornerstones of research on gaze control in reading (e.g., Altarriba, Kroll, Sholl, & Rayner, 1996; Calvo & Meseguer, 2002; Henderson & Ferreira, 1990; Henderson & Ferreira, 1993; Hyönä & Olson, 1995; Inhoff & Rayner, 1986; Just & Carpenter, 1980; Kennison & Clifton, 1995; Kliegl, 2007; Kliegl et al., 2004, 2006; Raney & Rayner, 1995; Rayner, 1977; Rayner & Duffy, 1986; Rayner & Fischer, 1996; Rayner et al., 2004).[8] The effect reflects the longer processing times associated with low-frequency words as compared to high-frequency words.

We will now propose an explanation for the pattern of current-word frequency effects observed in the present data. According to the foveal load hypothesis, a low-frequency word $n$ captures more attentional resources than a corresponding word of high-frequency. This should not only modulate the preview for the next word, but also reduce the additional time that is needed to process the low-frequency word. If the allocation of additional processing resources is strong enough (i.e., if the additionally captured resources are equal to the additional processing demands), this mechanism is capable of canceling out any immediacy effects of word frequency on fixation durations. In its most extreme version, a strong dynamical modulation of the perceptual span could even produce reversed, that is positive, effects of current-word frequency on fixation durations.

#### 4.2.3. Relative successor effect

According to the foveal load hypothesis, the parafovea-on-fovea frequency effect from word $n+1$ should depend on the frequency of the currently fixated word $n$. Previous studies did not find such an interaction (e.g., Henderson & Ferreira, 1990; Kliegl et al., 2006). We replicated this null effect for normal sentence reading. However, we observed a significant interaction for readers of shuffled text (see Fig. 4b for gaze durations and Fig. 5b for single fixation durations). This finding lends further support to the interpretation that the dynamical modulation was stronger in readers of shuffled text compared to readers of sentences.

#### 4.2.4. Effects of relative current-word frequency

The foveal load hypothesis further predicts that the strength of the current-word frequency effect depends on the length of the fixated word. The frequency effect should be stronger for long words than for short words, which is schematically illustrated in Fig. 1. For long words $n$ of high (Fig. 1a) or low (Fig. 1b) frequency, it is more or less only the currently fixated word $n$ that falls into the perceptual span. As a consequence, the effects of lexical processing are fully visible in the current-word frequency effect. Indeed, we found a strong frequency effect for long words in both reading conditions (Fig. 3a). The situation is different for short words. According to the foveal load hypothesis, a short low-frequency word is read with a narrowly focused perceptual span (Fig. 1d). In this case, all processing resources are focused on word $n$. If the currently fixated word $n$ is not only short but also high-frequent, the perceptual span should be enlarged, such that also the upcoming word $n+1$ falls into the span (Fig. 1c). Under the assumption of constant processing resources, this distribution of attention across two words can slow down the processing of the currently fixated word $n$, modulating the frequency effect observed for short words $n$. For normal reading, we found a small standard (i.e., negative) frequency effect for short words (Fig. 3a). For the shuffled text, this effect turned into a positive effect such that low-frequency words were actually fixated shorter than high-frequency words. Thus, the foveal load hypothesis is compatible with our experimental findings.

#### 4.2.5. Lag effects of relative word frequency

Another prediction that directly follows from such reasoning is that the preview for the upcoming word $n+1$ should depend on the interaction of word $n$ frequency and length. As noted above, long words $n$ fill more or less the whole perceptual span regardless of their frequency. As a consequence, preview for word $n+1$ will barely differ between conditions of low (Fig. 1a) and high (Fig. 1b) foveal load. Accordingly, word $n$ frequency should not strongly influence first fixation durations on the next word $n+1$. Indeed, we found weak effects of word $n$ frequency for readers of normal and for those of shuffled text if word $n$ was long (Fig. 3b). Again, the situation is different for short words $n$. During fixations on short low-frequency words $n$ the perceptual span is narrow and does not allow for much preprocessing of the next word (Fig. 1d). For short and high-frequency words $n$ the next word $n+1$ largely falls into the perceptual span (Fig. 1c). Strong parafoveal processing in this condition will reduce the processing needs for word $n+1$ when fixating on it. Thus, foveal load during fixations on short words should strongly influence the amount of parafoveal preprocessing. Empirically, the effect of word $n$ frequency on first fixation durations on word $n+1$ was strong for short words in both reading conditions, but stronger for readers of shuffled text (Fig. 3b). Thus, fleshing out the foveal load hypothesis within a parallel processing framework makes an interesting double-prediction concerning frequency effects of short words $n$: Word $n$ frequency should *weakly* influence fixation durations on word $n$ (or even show a reversed influence), but should *strongly* affect fixation durations on word $n+1$. Thus, there should be a trade-off between the two effects. The data support this prediction, as both effects are stronger for readers of shuffled compared to normal text.

#### 4.2.6. Regression probability

We examined how often readers regressed back to word $n$ after having fixated word $n+1$ once (and after having made a first-pass single fixation on word $n$) (Fig. 3c). Readers of normal text generated

[8] Going beyond fixation durations during reading, word frequency also affects word processing in many other psycho-linguistic tasks. That words, which occur frequently in a given language, are recognized more easily than words that appear less frequently is perhaps the single most robust finding in the whole literature on visual word recognition. The basic result holds across the entire range of laboratory tasks used to investigate reading. For example, frequency effects are seen in lexical decision [...], in naming [...], semantic classification [...], perceptual identification, [... and] spoken word recognition [...] and therefore appear to be a central feature of word recognition in general (Norris, 2006, p. 327; also see e.g., Monsell, 1991; Murray & Forster, 2004; Whaley, 1978).

significantly more regressions to low-frequency words. In contrast, when reading shuffled text more regressions were made to short words of high-frequency compared to short words of low-frequency. Thus, in shuffled-text reading, short high-frequency words did not only receive longer fixation times, but also more regressions than short low-frequency words.

Henderson and Ferreira (1990) demonstrated that the perceptual span is modulated by foveal load in normal reading. Here we applied this foveal load hypothesis to derive a qualitative model of our results on shuffled-text reading. It turned out that the foveal load hypothesis, combined with a processing gradient (PG) model of eye-movement control, provides a coherent theoretical for the explanation of a set of complicated and highly interacting effects. When reading difficult (i.e., low-frequency) words, shuffled text readers focus their attention so strong that they process these words even faster than easy (i.e., high-frequency) words. Likewise, processing of the next word is reduced. When fixating easy (high-frequency) words, on the other hand, readers of shuffled text widen their perceptual span such that high-frequency words – in particular if they are of short length – are fixated longer and attract more regressions compared to short words of low-frequency. At the same time parafoveal processing of word $n + 1$ is enhanced and fixation times on this word are reduced.

Why do we observe a stronger dynamical modulation of the perceptual span for readers of shuffled texts? As we speculated in the Introduction, readers of shuffled text may have focused on the processing of salient low-frequency content words when trying to remember the words in the shuffled text. In contrast, they may have widened their perceptual span when encountering high-frequency words because they did not expect to be probed about these words. Such processing would in fact be a good strategy because only content words, but not function words were queried in the memory task. It may be that readers were aware of this fact and adapted their processing to optimize the processing of task-relevant words. In sum, we propose that (a) a strong focus on low-frequency content words coupled with (b) limited processing resources that are spatially distributed via a dynamically modulated attentional gradient can lead to the disappearance or reversal of word-frequency effects during the reading of shuffled text.

### 4.3. Alternative explanations for changed frequency effects

As one of our findings, under certain conditions the effect of current-word frequency was strongly attenuated or even reversed when reading shuffled text. We argued that the dynamics of attention modulation in a PG model can qualitatively explain such an effect and the conditions under which it should occur. However, it could still be that frequency effects were reversed not because of the dynamics of attention modulation and eye-movement control but because high-frequency words were more difficult to process when reading shuffled text than low-frequency words. For example, short high-frequency words might slow down reading and attract regressions because they have more high-frequency orthographic neighbors, or because function words (as opposed to content words) are difficult to process when encountered in shuffled text. However, control analyses showed that these specific characteristics of short high-frequency words were not responsible for the observed patterns of results (see Online supplementary material).

Specific memorization processes related to the *mirror effect* (e.g., Reder et al., 2000) may provide another alternative explanation for why word-frequency effects were reversed. When studying a list of unrelated words, words of low-frequency were shown to be easier to recognize than words of high-frequency (e.g., Reder et al., 2000). It has repeatedly been shown that the effect is specific to retrieval and does not hold during encoding (e.g., de Zubicaray, McMahon, Eastburn, Finnigan, & Humphreys, 2005; Diana & Reder, 2006; Miozzo & Caramazza, 2003; Naveh-Benjamin, Craik, Guez, & Dori, 1998; Naveh-Benjamin & Guez, 2000; Rao & Proctor, 1984). However, it is possible that subjects are aware of their better recognition performance for low-frequency words. Thus, it could be that even though high-frequency words are more easily identified, readers actually invest more time in memorizing these words. This could potentially lead to reversed effects of word frequency because high-frequency words are fixated longer or because readers make more regressions to these words.

Finally, we hasten to emphasize that high-frequency words were not generally processed longer than low-frequency words. Effects of word frequency were often in the expected direction (see e.g., effects of successor and lag frequency). They were reversed only under very specific circumstances, in particular for short words. In addition, and critically, *reduced* or *reversed* effects of word $n$ frequency (on fixation durations on word $n$ and regression probability) were associated with an *enhancement* of these effects on fixation durations on word $n + 1$. Thus, a generally increased processing difficulty for high-frequency words cannot be responsible for the specific pattern of results in the present study.

### 4.4. PG and SAS models

We have shown that PG models incorporating the principles outlined above can, in principle, explain our results. A model like SWIFT might provide a parsimonious account based on a single mechanism, that is the modulation of the perceptual or attentional span by foveal load (see Engbert, 2007, for an implementation of the foveal load hypothesis with the SWIFT model). In contrast, given their basic principles, SAS models would not naturally predict the effects reported here. In particular, finding strong, experimental effects of distributed lexical processing and not finding the standard current-word frequency effect and, under some conditions, finding reversed current-word frequency effects, is not readily explained by the E-Z Reader model (Reichle, Warren et al., 2009).

As such, high-level effects of distributed processing provide a challenge for SAS models (Engbert & Kliegl, in press; Kliegl et al., 2006). These effects have been the subject of considerable debate (Rayner et al., 2007; but see Kliegl, 2007). As many as about 50 variables are known to influence word recognition (see Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). The corpus analyses by Kliegl and colleagues, finding pervasive effects of distributed processing, included only a limited number of such variables: They tested the effects of frequency, length, and predictability of words. If any of the remaining, uncontrolled variables (e.g., of the fixated word $n$) were correlated with the frequency of the next word $n + 1$, then corpus analyses could show significant successor effects of next-word frequency. However, these effects would, in fact, not stem from the processing of the next word $n + 1$, but instead from lexical processing of the currently fixated word $n$ (cf., Rayner et al., 2007). Rayner and colleagues (2007) implemented this hypothesis to simulate results from Kliegl and colleagues (2006) with the E-Z Reader model. They assumed that the predictability of word $n + 1$ was correlated with an unobserved variable influencing lexical processing of word $n$. Introducing this simple correlation was sufficient for the E-Z Reader model to show substantial effects of word $n + 1$ predictability on fixation durations on word $n$. Introducing a similar correlation with word $n + 1$ frequency would enable the E-Z Reader model to show substantial effects of word $n + 1$ frequency on fixation durations on word $n$. As noted above (see Section 4.1), correlations between neighboring word properties are absent in shuffled text. Each word was randomly selected for each shuffled word list and each reader separately. Therefore, unobserved properties of the fixated word $n$ cannot be systematically related to the frequencies of neighboring words. Thus, our results on distributed lexical processing are of experimental nature. They impose boundary conditions for computational models of reading.

According to proponents of the E-Z Reader model, simulations could in principle accommodate lexical influences from neighboring words if these were due to mislocated fixations (e.g., Rayner et al., 2007). In reading, due to oculomotor error in saccade programming a significant proportion of fixations are mislocated in that they fall on words to the left or right of the intended target word (Engbert & Nuthmann, 2008). In E-Z Reader, it is the intended rather than the fixated word that will receive lexical processing. However, we believe that numerical simulations are necessary to explore the possibility that mislocated fixations can induce parafoveal-on-foveal effects in SAS models. Some empirical evidence for the mislocation hypothesis has been reported (Drieghe, Rayner, & Pollatsek, 2008; subsequently challenged by Kennedy (2008)), but the results were not substantiated by quantitative estimates. Moreover, mislocated fixations trigger short-latency saccades to produce the fixation-duration IOVP effect (Nuthmann et al., 2005). Thus, these short fixations are not triggered by lexical word processing, and should be independent of the frequency of the intended (neighboring) word. This, however, is inconsistent with Rayner et al.'s (2007) hypothesis that mislocated fixations cause effects of the neighboring (intended) word frequency on fixation durations on the fixated word *n* (cf., Kennedy, 2008).

In general, drawing conclusions from the shuffled-text reading task about theoretical models of eye-movement control is preliminary. First, numerical simulations of the models need to be carried out. Second, it is unclear at present how different cognitive processes (e.g., related to memory demands) influence eye movements during reading of shuffled text compared to normal text reading. Therefore, further empirical as well as computational research is needed to illuminate these issues.

## 5. Conclusion

In the present paper we introduced the shuffled-text reading paradigm as a new paradigm to study the interactive control of eye movements by higher-level cognitive and lower-level visuomotor factors. We found that a number of variables known

**Table A1**
Eye-movement statistics for reading shuffled and normal text.

| Variable | shuffled text | | normal text | | t-test | | |
|---|---|---|---|---|---|---|---|
| | *M* | (*SD*) | *M* | (*SD*) | *t* | *df* | *P* |
| N of readers | 30 | | 30 | | | | |
| N of fixations/sentence | 10.1 | (2.0) | 7.8 | (1.4) | 5.14 | 51 | < 0.001 |
| N of sentences/reader | 139 | (7) | **137** | (8) | | | |
| *Fixation probabilities* | | | | | | | |
| skipping (p0) | 0.10 | (0.06) | 0.21 | (0.08) | -6.45 | 55 | < 0.001 |
| single fixation (p1) | 0.70 | (0.06) | 0.67 | (0.06) | 1.69 | 58 | 0.10 |
| double-plus fixation (p2+) | 0.16 | (0.07) | 0.08 | (0.04) | 5.41 | 48 | < 0.001 |
| regression (prg) | **0.06** | (0.04) | **0.06** | (0.04) | | | |
| mean saccade length (letters) | 6.1 | (0.9) | 7.6 | (1.2) | -5.85 | *55* | < 0.001 |
| *Fixation position (letter)* | | | | | | | |
| single fixation (l0) | 2.5 | (0.3) | 2.7 | (0.2) | -3.22 | 53 | < 0.01 |
| 1st of multiple (l1) | **2.0** | (0.4) | **2.2** | (0.6) | | | |
| 2nd of multiple (l2) | **5.5** | (0.4) | **5.5** | (0.8) | | | |
| *Fixation duration (ms)* | | | | | | | |
| single fixation (d0) | 254 | (40) | 213 | (32) | 4.37 | 55 | < 0.001 |
| 1st of multiple (d1) | 227 | (32) | 199 | (30) | 3.50 | 58 | < 0.001 |
| 2nd of multiple (d2) | 197 | (36) | 172 | (36) | 2.74 | 58 | < 0.01 |
| gaze duration | 293 | (58) | 231 | (37) | 4.89 | 50 | < 0.001 |
| Reading rate (words/min) | 193 | (47) | 250 | (46) | -4.74 | 58 | < 0.001 |

*Note*. No invalid fixations were removed. Data are from right eye. Mean n of fixations (N), regression probability (prg) and reading rate are based on all fixations; all other measures are based on first-pass reading. Welch t-tests over participants were used to test differences between normal and shuffled text reading. Values of **non**-significant differences (*ps* > .25) are printed in **bold**.

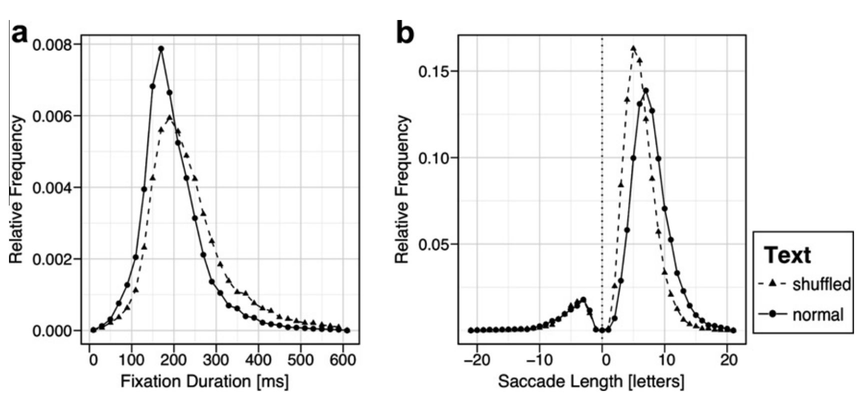


**Fig. A1.** Global analyses. (a) Distribution of all observed valid fixation durations during reading of randomly shuffled text (triangles and dashed line) vs. normal reading (circles and solid line). Displays the corresponding mean frequency distributions. Relative proportions of fixation durations are displayed for 31 levels (from 0 ms up to 620 ms in 20-ms steps). (b) Distributions of all observed saccade lengths. Negative saccade lengths indicate regressive saccades.

**Table B1**
Results from linear mixed models fit by restricted maximum likelihood (REML): Means, standard errors, and *t*-values of fixed effects on fixation durations; variances and standard deviations of the random effects.

| | Log gaze durations | | | Log single fixation durations | | |
|---|---|---|---|---|---|---|
| *Fixed effects* | | | | | | |
| | Estimate | SE | *t*-Value | Estimate | SE | *t*-Value |
| Intercept | 5.540 | 0.029 | 191.3 | 5.489 | 0.030 | 182.2 |
| *Word n* | | | | | | |
| Frequency (frq) | −**0.0002** | **0.006** | −**0.04** | *0.009* | *0.005* | *1.77* |
| frq * frq | 0.020[a] | 0.005 | 4.3 | 0.017[a] | 0.004 | 4.5 |
| 1/length (lgth) | −0.739 | 0.072 | −10.2 | 0.265 | 0.063 | 4.2 |
| *Word n − 1* | | | | | | |
| Frequency | −0.035 | 0.002 | −15.1 | −0.034 | 0.003 | −12.5 |
| 1/length | 0.247[a] | 0.026 | 9.5 | 0.207 | 0.035 | 5.9 |
| *Word n + 1* | | | | | | |
| Frequency | −0.016 | 0.002 | −6.9 | −0.016 | 0.002 | −6.6 |
| 1/length | 0.119[a] | 0.025 | 4.8 | 0.114[a] | 0.026 | 4.4 |
| *Viewing position* | | | | | | |
| Last sacc. amplit. | 0.017 | 0.001 | 13.6 | 0.027[a] | 0.001 | 30.1 |
| pos in word | −0.138 | 0.017 | −8.3 | −0.082[a] | 0.013 | −6.2 |
| pos * pos | −1.088 | 0.050 | −21.6 | −0.348[a] | 0.038 | −9.3 |
| Next sacc. amplit. | −0.007 | 0.001 | −5.6 | 0.011[a] | 0.001 | 10.4 |
| *Interactions* | | | | | | |
| (frq *n*)/(lgth *n*) | 0.401[a] | 0.055 | 7.3 | −**0.063** | **0.048** | −**1.3** |
| (frq *n*) * (frq *n* − 1) | 0.015 | 0.001 | 10.4 | 0.019 | 0.002 | 12.3 |
| (frq *n*) * (frq *n* + 1) | 0.010 | 0.002 | 5.5 | 0.009 | 0.002 | 5.1 |
| (frq *n* + 1)/(lgth *n*) | 0.060 | 0.025 | 2.4 | 0.099[a] | 0.021 | 4.7 |
| Slope-differences between shuffled and normal PSC reading | | | | | | |
| Experim. cond. (Exp) | −0.296 | 0.040 | −7.4 | −0.270 | 0.042 | −6.4 |
| *Word n* | | | | | | |
| Exp * frq | −0.026 | 0.004 | −6.7 | −0.028 | 0.004 | −7.0 |
| Exp * frq * frq | [a] | | | [a] | | |
| Exp * lgth | 0.483 | 0.054 | 9.0 | −**0.009** | **0.057** | −**0.2** |
| *Word n − 1* | | | | | | |
| Exp * frq | −**0.004** | **0.003** | −**1.2** | *−0.008* | *0.004* | *−1.7* |
| Exp * lgth | [a] | | | *0.105* | *0.054* | *1.94* |
| *Word n + 1* | | | | | | |
| Exp * frq | *0.006* | *0.003* | *1.86* | **0.005** | **0.003** | **1.6** |
| Exp * lgth | [a] | | | [a] | | |
| *Viewing position* | | | | | | |
| Exp * last sacc. amp. | 0.007 | 0.002 | 4.0 | [a] | | |
| Exp * pos in word | **0.033** | **0.023** | **1.5** | [a] | | |
| Exp * pos * pos | 0.453 | 0.070 | 6.5 | [a] | | |
| Exp * next sac. amp. | 0.013 | 0.002 | 7.8 | [a] | | |
| *Interactions* | | | | | | |
| Exp * (frq *n*)/(lgth *n*) | [a] | | | 0.177 | 0.037 | 4.8 |
| Exp * (frq *n*) * (frq *n* − 1) | −0.019 | 0.002 | −7.9 | −0.018 | 0.002 | −7.2 |
| Exp * (frq *n*) * (frq *n* + 1) | −0.010 | 0.003 | −3.3 | −0.010 | 0.002 | −4.0 |
| Exp * (frq *n* + 1)/(lgth *n*) | 0.102 | 0.043 | 2.4 | [a] | | |

| | | Log gaze durations | | Log single fixation durations | |
|---|---|---|---|---|---|
| *Random effects* | | | | | |
| Groups | Name | Variance | Std. Dev. | Variance | Std. Dev. |
| Word ID | Intercept | 0.0085 | 0.092 | 0.0045 | 0.067 |
| Reader | Intercept | 0.0237 | 0.153 | 0.0262 | 0.162 |
| Residual | | 0.1232 | 0.351 | 0.0839 | 0.290 |
| *N* of fixations | | 38,738 | | 24,433 | |
| AIC | | 30,282 | | 9,921 | |
| BIC | | 30,548 | | 10,148 | |
| logLik | | −15,110 | | −4,933 | |

*Note*: All data are from right eye (60 readers; 550 unique word IDs). **Non**-significant coefficients are set in bold ($t < 1.645$). *Marginally* significant coefficients are set in italics ($1.645 \leqslant t < 1.96$). Shuffled text reading is the reference condition. Experimental condition (Exp) depicts the contrast between that reference condition and the normal text reading condition using a dummy-coded factor.
[a] The slope-difference between shuffled and normal text reading was not significant for these effects, thus the interactions of the respective effect with experimental condition was dropped from the model. The main effect reflects the average effect in shuffled and normal reading.

to influence eye movements in reading showed similar effects when reading shuffled texts. Thus, the basic mechanisms of visuo-motor and lexical processing are at work independent of whether meaningful sentences are presented or not. However, shuffled text has an impact on global parameter settings and modulates strategies for information processing in reading. We demonstrated two such influences. First, our findings add to the body of literature suggesting that the predictability of words eases their processing and speeds up reading (e.g., Balota et al., 1985; Ehrlich & Rayner, 1981), albeit from a novel perspective. In the shuffled-text reading paradigm, word predictability is removed while word frequency remains intact. We showed that this manipulation of word predictability as well as potential differences in the memorization of words slowed down reading. The findings also contribute to the current debate about serial as opposed to parallel processing of words in a sentence (Engbert & Kliegl, in press; Reichle, Liversedge et al., 2009). We observed distinct experimental effects of spatially distributed processing (Kliegl et al., 2006), indicating that several words are simultaneously affecting fixation duration at a time. These effects were more strongly modulated by foveal load in the shuffled reading task as compared to normal reading.

## Acknowledgments

This work was supported by Deutsche Forschungsgemeinschaft (Grant EN 471/2-1). We would like to thank Reinhold Kliegl for his statistical advice, Sarah Risse for her support during data analysis and the writing process, and Keith Rayner for valuable comments on previous versions of this manuscript. We also thank Denis Drieghe and other anonymous reviewers for their helpful comments.

## Appendix A. Global analyses

See Table A1 and Fig. A1.

## Appendix B. LME models

See Table B1.

## Appendix C. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2010.08.005.

## References

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition, 24*(4), 477–492.

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

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*(4), 390–412.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575–589.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. P. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General, 133*, 283–316.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*(3), 364–390.

Bates, D., & Sakar, D. (2008). *lme4: Linear mixed-effect models using S4 classes*, Version 0.999375-15.

Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical, and contextual factors. *The Spanish Journal of Psychology, 5*(1), 66–77.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42–45.

de Zubicaray, G. I., McMahon, K. L., Eastburn, M. M., Finnigan, S., & Humphreys, M. S. (2005). fMRI evidence of word frequency and strength effects in recognition memory. *Cognitive Brain Research, 24*(3), 587–598.

Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 805–815.

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

Ehrlich, S. E., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior, 20*, 641–655.

Engbert, R. (2007). *Reading with a dynamic processing span. Paper presented at the 14th European conference on eye movements, ECEM 14*, University of Potsdam.

Engbert, R., & Kliegl, R. (in press). Parallel graded attention models of reading. In: S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), Oxford handbook on eye movements. Oxford: Oxford University Press.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43*(9), 1035–1045.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42*(5), 621–636.

Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences of the United States of America, 103*(18), 7192–7197.

Engbert, R., & Nuthmann, A. (2008). Self-consistent estimation of mislocated fixations during reading. *PLoS ONE, 3*(e1534), 1–6. doi:10.1371/journal.pone.0001534.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*(4), 777–813.

Geyken, A. (2006). The DWDS-corpus: A reference corpus for the German language of the 20th century. In C. Fellbaum (Ed.), *Collocations and idioms: Linguistic, lexicographic, and computational aspects*. London: Continuum 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*(3), 417–429.

Henderson, J. M., & Ferreira, F. (1993). Eye-movement control during reading – Fixation measures reflect foveal but not parafoveal processing difficulty. *Canadian Journal of Experimental Psychology – Revue Canadienne De Psychologie Experimentale, 47*(2), 201–221.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(6), 1430–1440.

Inhoff, A. W., Eiter, B. M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 979–995.

Inhoff, A. W., Pollatsek, A., Posner, M. I., & Rayner, K. (1989). Covert attention and eye movements during reading. *Quarterly Journal of Experimental Psychology Section A – Human Experimental Psychology, 41*(1), 63–89.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading – Effects of word frequency. *Perception & Psychophysics, 40*(6), 431–439.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading – From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Kennedy, A. (2008). Parafoveal-on-foveal effects are not an artifact of mislocated saccades. *Journal of Eye Movement Research, 2*(1), 2. 1–10.

Kennedy, A., & Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research, 45*(2), 153–168.

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

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

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition, 18*(5), 655–681.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology, 16*(1–2), 262–284.

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

Kliegl, R., Risse, S., & Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word $n + 2$. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1250–1255.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences, 4*(1), 6–14.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203.

Mason, R. A., & Just, M. A. (2004). How the brain processes causal inferences in text – A theoretical account of generation and integration component processes utilizing both cerebral hemispheres. *Psychological Science, 15*(1), 1–7.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics, 17*(6), 578–586.

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture–word interference paradigm. *Journal of Experimental Psychology: General, 132*(2), 228–252.

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

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

Morton, J. (1964). The effects of context upon speed of reading, eye movements and eye-voice span. *Quarterly Journal of Experimental Psychology, 16*(4), 340–354.

Murray, W. S., & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review, 111*(3), 721–756.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language, 26*(4), 453–465.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology: Learning Memory and Cognition, 24*(5), 1091–1104.

Naveh-Benjamin, M., & Guez, J. (2000). Effects of divided attention on encoding and retrieval processes: Assessment of attentional costs and a componential analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1461–1482.

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal Bayesian decision process. *Psychological Review, 113*(2), 327–357.

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research, 49*(3), 322–336.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research, 45*(17), 2201–2217.

Nuthmann, A., Engbert, R., & Kliegl, R. (2007). The IOVP effect in mindless reading: Experiment and modeling. *Vision Research, 47*(7), 990–1002.

Nuthmann, A., & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision, 9*(5), 31. 1–28, doi:10.1167/9.5.31.

Pinheiro, J., & Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.

Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye- movement control from the perceptual span in reading. *Perception and Psychophysics, 40*(2), 123–130.

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

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006a). Serial processing is consistent with the time course of linguistic information extraction from consecutive words during eye fixations in reading: A response to Inhoff, Eiter, and Radach (2005). *Journal of Experimental Psychology: Human Perception & Performance, 32*(6), 1485–1489.

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

R Development Core Team (2008). *R: A language and environment for statistical computing*, Version 2.6.2, R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.or>g.

Radach, R., & Heller, D. (2000). Relations between spatial and temporal aspects of eye movement control. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 165–191). Amsterdam: Elsevier Science.

Raney, G. E., & Rayner, K. (1995). Word frequency effects and eye movements during two readings of a text. *Canadian Journal of Experimental Psychology – Revue Canadienne De Psychologie Experimentale, 49*(2), 151–173.

Rao, K. V., & Proctor, R. W. (1984). Study-phase processing and the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*(3), 386–394.

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition, 5*(4), 443–448.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*(8), 1457–1506.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. D. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 720–732.

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

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58*(5), 734–747.

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and Performance XII: The psychology of reading* (pp. 327–362). London: Erlbaum.

Rayner, K., Pollatsek, A., Drieghe, D., Slattery, T. J., & Reichle, E. D. (2007). Tracking the mind during reading via eye movements: Comments on Kliegl, Nuthmann, and Engbert (2006). *Journal of Experimental Psychology: General, 136*(3), 520–529.

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin &Review, 3*, 238–244.

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

Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S. P. (2003). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier.

Reder, L. M., Nhouyvanisvong, A., Schunn, C. D., Ayers, M. S., Angstadt, P., & Hiraki, K. (2000). A mechanistic account of the mirror effect for word frequency: A computational model of remember–know judgments in a continuous recognition paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(2), 294–320.

Reichle, E. D., Liversedge, S. P., Pollatsek, A., & Rayner, K. (2009). Encoding multiple words simultaneously in reading is implausible. *Trends in Cognitive Sciences, 13*(3), 115–119.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 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*(1), 4–22.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*(4), 445–526.

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

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

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research, 7*(1), 34–55.

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science, 17*(9), 742–746.

Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *Quarterly Journal of Experimental Psychology Section A – Human Experimental Psychology, 52*(4), 1021–1046.

Starr, M. S., & Rayner, K. (2001). Eye movements during reading: Some current controversies. *Trends in Cognitive Sciences, 5*(4), 156–163.

Vitu, F., Brysbaert, M., & Lancelin, D. (2004). A test of parafoveal-on-foveal effects with pairs of orthographically related words. *European Journal of Cognitive Psychology, 16*(1–2), 154–177.

Vitu, F., Lancelin, D., & d'Unienville, V. M. (2007). A perceptual-economy account for the inverted-optimal viewing position effect. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1220–1249.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 41*(25–26), 3513–3533.

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading – Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics, 57*(3), 352–364.

Whaley, C. P. (1978). Word–nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17*(2), 143–154.

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*(5), 891–896.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer.