Reading strategy modulates parafoveal-on-foveal effects in sentence reading

Christiane Wotschack & Reinhold Kliegl

University of Potsdam, Germany

Key words: Fixational selectivity, Parafoveal-on-foveal effects, Reading strategy, Distributed processing, Eye movements in reading

Address for correspondence:

Christiane Wotschack

University of Potsdam

Department of Psychology

Karl-Liebknecht-Str. 24/ 25

14476 Potsdam, Germany

Fon: +49 331 977 2869, Fax: +49 331 977 2793

Email: c.wotschack@fu-berlin.de

Abstract

Task demands and individual differences have been linked reliably to word skipping during reading. Such differences in fixation probability may imply a selection effect for multivariate analyses of eye-movement corpora if selection effects correlate with word properties of skipped words. For example, with fewer fixations on short and highly frequent words the power to detect parafoveal-on-foveal effects is reduced. We demonstrate that increasing the fixation probability on function words with a manipulation of the expected difficulty and frequency of questions reduces an age difference in skipping probability (i.e., old adults become comparable to young adults) and helps to uncover significant parafoveal-on-foveal effects in this group of old adults. We discuss implications for the comparison of results of eye-movement research based on multivariate analysis of corpus data with those from display-contingent manipulations of target words.

Reading strategy modulates parafoveal-on-foveal effects in sentence reading

Reading research has (at least) two traditions. Originally, reading research focused on individual and developmental differences and evaluated the diagnostic power of global eye-movement parameters such as skipping probability or average fixation duration (Buswell, 1922; Huey, 1908; for a recent update see Radach & Kennedy, 2004). In this line of research, the emphasis has been on differences in reading strategies that may prevail for an entire sentence or text. Starting with McConkie and Rayner's (1975) use of gaze-contingent display changes, a second line of research commenced that focused on disentangling perceptual, oculomotor, and psycholinguistic effects on fixation durations. In this line of research, the emphasis has been on the perceptual span, a critical region of three to four words in a sentence around the point of fixation. For skilled readers the perceptual span extends about 3-4 characters to the left and about 14-15 characters to the right of fixation (McConkie & Rayner, 1975; Rayner, 1975).

In our study we draw attention to an important implication of "higher-level" control parameters of global reading strategies, such as depth of reading comprehension, for the perceptual-span perspective and in particular for the detection of effects of properties of upcoming words that have not been fixated yet (i.e., parafoveal-on-foveal effects, Kennedy, 2000). We take as a starting point that differences in reading strategy may lead to differences in skipping probabilities and that these differences in turn may lead to differences in detecting parafoveal-on-foveal effects. Our results may help us understand differences between results from multivariate analyses of fixation durations (e.g., Kennedy & Pynte, 2005; Kliegl, Nuthmann, & Engbert, 2006) and results obtained with fixation durations linked to experimental manipulations of target words (see Kliegl, 2007; Rayner, Pollatsek, Drieghe, Slattery, & Reichle, 2007, for a discussion).

Reading is a selective process and about 10-30% of the words are skipped during reading. Short, high frequent, and high predictable words are skipped more often than longer words of lower frequency or predictability (Drieghe, Desmet, & Brysbaert, 2007; Drieghe, Pollatsek, Staub, & Rayner, 2008; Rayner, 1998). Therefore, function words, such as determiners, prepositions or conjunctions, are prime candidates for skipping because they are usually short and very frequent in language. In other words, a lower skipping rate is usually associated with a higher selection of fixated function words.

On top of these general selection effects, readers vary in their individual eye movement patterns during reading. For example, older adults generally show a higher skipping rate in first-pass reading and more regressive eye movements than young adults (Kliegl, Grabner, Rolfs, & Engbert, 2004; Laubrock, Kliegl, & Engbert, 2006; Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006). Rayner et al. suggested that older adults may engage in a more risky reading strategy than young adults. Of course, a global parameter such as skipping probability is also susceptible to experimental manipulations of task demand. For example, skipping probability will be reduced if very careful reading is induced (e.g., Tinker, 1958). In our study, we used easy and difficult questions to manipulate the degree of careful reading (and skipping probability) in different groups of young and older adults.

Research on eye movement control in reading has consistently demonstrated that characteristics of a fixated word, such as word length, word frequency, or word predictability, modulate fixation duration (for a review see Rayner, 1998). For example, single fixation duration and gaze duration on low frequency words are shorter than on high frequency words of the same length (Just & Carpenter, 1980; Rayner & Duffy, 1986). Importantly, under the assumption of distributed processing within the perceptual span, both frequency and predictability of the word to the right of fixation (n+1) may also affect processing time of the foveal word n. Such effects are referred to as a parafoveal-on-foveal effect (Kennedy, 2000; Kennedy, Pynte, & Ducrot, 2002).

Several studies demonstrated lexical and sublexical parafoveal-on-foveal effects on fixation duration, providing evidence from experimentally manipulated target words as well as from multivariate analyses of eye-movement corpora. In an experiment, in which participants read a sequence of five content words and searched for a word of a certain semantic category, Kennedy et al. found a parafoveal-on-foveal effect of word informativeness and frequency of the upcoming word on gaze duration (Kennedy, 2000; Kennedy, et al., 2002). In text reading, Kennedy and Pynte (2005) demonstrated frequency effects of word n+1 on gaze duration and on single-fixation duration on the fixated word. Importantly, the effect of parafoveal word frequency on fixation duration was only reported for short foveal words. The authors argued that parafoveal-on-foveal effects vary systematically as a function of the length of the foveal and parafoveal word involved (Kennedy & Pynte, 2005).

Kliegl, Nuthmann and Engbert (2006) also reported a significant negative effect of the frequency of the upcoming word n+1 on single fixation duration of the fixated word n in normal sentence reading (i.e., shorter fixation durations on word n with high-frequency words n+1). Comparable to the results from Kennedy and Pynte, they demonstrated an interaction of foveal word length with frequency of the parafoveal word. Kliegl et al.

(2006) argued that if word n is short there is room for preprocessing word n+1 within the perceptual span. Moreover, despite the significant positive correlation of frequency and predictability and the negative effect of frequency of the upcoming word, predictability of word n+1 had a significant positive effect on single-fixation duration on words in these analyses; that is, single fixations on word n were longer for highly predictable words n+1. These results were stable across nine different samples of readers. They were also obtained with a linear mixed model analysis, including also skipping status of word n-1 and word n+1 as well as the syntactic category (content vs. function word) of the fixated word and its two neighbours (Kliegl, 2007). In summary, multivariate analyses of single-fixation durations demonstrated reliable parafoveal-on-foveal frequency and predictability effects under simultaneous statistical control of a large number of correlated predictors.

Recently, Schad, Nuthmann, and Engbert (2010) also supported the reliability of parafoveal-on-foveal effects in sentence reading using a multivariate analysis. They demonstrated that the frequency of the parafoveal word had a reliable effect on gaze and single-fixation duration during reading of shuffled text. Importantly, in a randomly shuffled text the neighbouring words are not systematically correlated, a critique that has been raised against parafoveal-on-foveal effects based on multivariate statistics (Rayner, et al., 2007).

From a processing perspective, the size of the perceptual span as well as the length of words in the current perceptual span are critical for the emergence of parafoveal-on-foveal effects. Moreover, the asymmetric shape of the perceptual span (McConkie & Rayner, 1975) appears to be subject to modulation by foveal load (Henderson & Ferreira, 1990, 1993; see also Rayner & Pollatsek, 1987). These studies provided some evidence that the perceptual span responds to the experimental manipulation of foveal processing load; in other words, preprocessing of the upcoming word n+1 is more efficient when word n is easy to process. In our study, we focus on the interaction between properties of the fixated word n and the upcoming word n+1 in multivariate analyses of single-fixation durations, taking into account not only word length but also word frequency and syntactic category of words in the perceptual span.

Though the current empirical evidence supports the reliability of parafoveal-on-foveal effects, explanations of these effects have been critically discussed and the theoretical controversies are far from resolved (Kliegl, 2007; Miellet, O'Donnell, & Sereno, 2009; Rayner, et al., 2007; Reichle, Liversedge, Pollatsek, & Rayner, 2009). Importantly, many more short words are typically included in multivariate analyses on eye-movement data of normal sentence reading than in analyses of fixation durations on target words in experimental designs (Kennedy, 2008). Therefore, chances were probably better to detect such effects in multivariate statistical analyses than in the analyses of target words, which are usually long content words.

We propose that it is not only the composition of the reading material per se that leads to differences in effects of parafoveal word properties on foveal processing, but that different reading strategies are likely to generate differences in the composition of the data base. For example, if old adults skip words more frequently than young adults during first-pass reading, they presumably contribute fewer short words than young adults to the multivariate analysis. Given this difference in the composition of the data base, one may ask whether such selection effects translate into age differences in parafoveal-on-foveal effects. Specifically, we hypothesize that differences in the probability of fixating short function words will have an impact on the size of parafoveal-on-foveal frequency effects, simply because with fewer fixations on functions words, there is a reduced statistical power to detect these effects.

Careful reading increases the number of fixations (Heller, 1985) and presumably also the fixation probability for short function words. Therefore, we had different groups of young and older adults read the same sentences under two experimental conditions that induced different reading styles by manipulating the difficulty and the frequency of questions after the sentences.

In summary, the goal of the study was to answer two questions: (1) Does the adaption of reading style to question difficulty affect fixational selectivity? (2) Is the selectivity of the fixated words related to parafoveal-on-foveal effects? We focused on selection effects in single-fixation cases and evaluate frequency effects of the fixated word n (foveal) and the upcoming word n+1 (parafoveal).

Method

Participants

Data of four groups are compared in this study. A group of 24 high school students and a group of 32 older readers read the Potsdam sentence corpus (PSC; Kliegl et al., 2004; 2006) with easy questions[1]. The young adults averaged 17.6 years of age (SD = 0.6, range: 16 - 18 years) and the older adults averaged 70.6 years of age (SD = 4.0, range: 65 - 84 years). An age matched group of 30 high school students and 25 old readers read the PSC with frequent, difficult questions. The hard-question young adults averaged 18.5 years of age (SD = 0.9, range = 17 – 20 years) and the hard-question old adults averaged 68.0 years of age (SD = 3.3, range = 65 – 76 years). Data of two participants from the hard-question old group were excluded from analysis because subjects provided less than 150 fixations to the whole data pool. All participants were native speakers of German. They all had normal or corrected to normal vision.

Comparisons between the easy-question and hard-question (same-age) groups revealed no significant differences in age, in scores on Lehrl's (1977) multiple-choice measure of vocabulary, or in Wechsler's (1964) Digit-Symbol-Test (all $ps > 0.05$). The age groups exhibited the usual psychometric profile of cognitive-aging research. Old adults scored significantly higher in vocabulary, $F(1,105) = 46.23$, $MSe = 4.73$, $p < 0.001$, and significantly lower in digit symbol substitution, $F(1,105) = 52.79$, $MSe = 90.8$, $p < 0.001$, than young readers.

Sentence and question material

The PSC comprises 144 German single sentences (1,138 words), which represent a large variety of grammatical structure. Sentences range from 5 to 11 words (M = 7.9, SD = 1.4). Norms on various psycholinguistic variables such as word length, word frequency (Geyken, 2007; Heister, et al., 2010) and predictability norms from an independent study are given for each word in the PSC (see Kliegl et al., 2004, for additional details about the sentence material).

In the easy-question condition, easy multiple-choice comprehension questions were asked after 27% of the sentence trials. Questions used identical wording of the preceding sentence and three alternative choices were provided. These questions were similar to single-word probe tasks and therefore correct responses were often possible solely by visual word recognition of the answering options.

In the hard-question condition, a difficult three alternative multiple-choice comprehension question was asked after each sentence of the PSC. The combination of questions and the alternative choices were designed to reduce the verbatim overlap with the original sentence in order to make a purely visual solution of the question impossible (e.g., by a simple word form recognition). The content of all questions aimed at testing a complete propositional representation of the sentence, thereby inducing a very deliberate reading strategy while keeping the sentence material identical.

Apparatus

Single sentences of the PSC were presented on the center-line of a 21-in. EYE-Q 650 Monitor (832 pixels x 632 pixels resolution; frame rate 75 Hz; font: regular, New Courier, 12 point). Participants were seated at a 60cm-distance in front of the monitor with the head positioned on a chin rest. One letter subtended 0.38° of visual angle. Eye movements of three samples were recorded with an EyeLink-II system (SR Research Ltd, Osgoode, ON, Canada) with 500-Hz-sampling rate. The easy-question group of old adults was recorded with an EyeLink-I system with a 250-Hz-sampling rate. All recordings and calibrations were binocular.

Procedure

In both reading conditions, participants were calibrated with a standard nine-point grid for both eyes. They were instructed to read the sentence for comprehension and to fixate on a dot in the lower right corner of the monitor to signal the completion of a trial. After validation of calibration accuracy, a fixation dot appeared on the left side of the center-line on the monitor. If the eye tracker identified a fixation on the fixation spot, a sentence was presented so that the midpoint between the beginning and the center of the first word was positioned at the location of the fixation spot. Therefore, each sentence-initial word was read from a word-specific optimal viewing position (O'Regan & Levy-Schoen, 1987). Sentences were shown until participants looked at the lower right corner of the screen. Then in 27% (easy-question condition) or 100% (hard-question condition), respectively, the sentence was replaced by a three-alternative multiple-choice question the participant answered via a mouse click. After every 15 sentences, a complete recalibration with the nine-point grid was presented. Ten training trials preceded 144 experimental trials. For their participation, subjects either received course credit or were paid 5-7 €/ hour.

Data selection

Saccades were detected using a velocity-based algorithm introduced by Engbert and Kliegl (2003; see also Engbert & Mergenthaler, 2006). While reading saccades, and thus fixations, were detected as binocular events, all statistical analyses were based only on right eye data. All detected fixations were included in word based summary statistics. First and last fixations in a sentence and fixations on the first and last words were excluded from the analysis of selection effects and from linear mixed models.

Linear mixed models (LMMs)

The effects of various individual-based, lexical as well as oculomotor predictors on log single fixation duration (SFD) were evaluated in a single sweep with linear mixed models (LMMs; Pinheiro & Bates, 2000), taking as a reference a repeated-measures multiple regression model based on 222 readers reading the PSC (Kliegl, et al., 2006).

The predictors included three random factors (subject ID, word ID, sentence ID) and 24 fixed effects plus interaction terms.

Reader-level predictors included trial number, vocabulary score, and reading condition (easy-, hard-question). To further account for individual differences in reading style without increasing model complexity, the 1st and 2nd component of a principal component analysis (PCA) based on ten relevant subject-level predictors were included[2].

Word-level predictors included incoming and outgoing saccade amplitude, log word frequency (linear, quadratic, and cubic trend), logit word predictability, word length (using the reciprocal value 1/length), the linear and quadratic components of relative fixation position (defined as letter-position/word-length scaled to zero, representing the center of the word).

For the critical tests of effects indicating distributed processing, log word frequency, logit predictability, and word length (reciprocal value) of the previous word n-1 as well as log word frequency, logit predictability, and word length (reciprocal value) of the upcoming word n+1 were also included as fixed effects (for further details on model fitting see Wotschack, 2009). Since we were interested in effects of reading condition and to reduce model complexity, models were built separately for young and old readers.

Analyses were carried out with the *lmer* function of the *lme4* package (Bates & Maechler, 2009) in the R environment for statistical computing (R Development Core Team, 2009). All continuous covariates were centered on the respective subjects' means.

Results

Condition and age effects

An ANOVA of response accuracy of the multiple-choice comprehension yielded a significant main effect of reading condition, $F(1,105) = 41.1$, $MSe = 0.0008$, $p < .001$, and a significant interaction between age and reading condition, $F(1,105) = 11.43$, $MSe = 0.0008$, $p < .01$. Though response accuracy was still at a high level, both young and old readers were less accurate in the hard-question (young: 95%, old: 92% correct) than in the easy-question condition (young: 97%, old: 97% correct); old readers' accuracy was more affected by question difficulty. An ANOVA of response times showed that old readers took more time to answer the questions than young adults, $F(1,105) = 79$, $MSe = 0.03$, $p < .001$.

Summary statistics and selection effects

Word-based summary statistics for each group are listed in Table 1. Separate ANOVAs for each age group revealed several main effects of reading condition. The hard-question young group produced significantly fewer first-pass single fixation cases, $F(1,52) = 7.1$, $MSe = 0.004$, $p < .05$, but more regressions, $F(1,52) = 35$, $MSe = 0.004$, $p < .001$, than the easy-question young group. In line with an increased proportion of second-pass reading, total reading time was longer in hard-question young than in easy-question young, $F(1,52) = 18.3$, $MSe = 4216$, $p < .001$.

(Table 1 about here)

For old readers, effects of reading condition were obtained for fixation probability, fixation duration, and fixation position. In first pass reading, the hard-question old group had a higher probability of fixating a word three or more times, $F(1, 53) = 9.7$, $MSe = 0.002$, $p < 0.01$, and made significantly more regressions, $F(1, 53) = 8.9$, $MSe = 0.02$, $p < .001$. The difference between conditions in the proportion of double-fixation cases was marginally significant, $F(1, 53) = 3.7$, $MSe = 0.002$, $p = .059$. Importantly, the easy-question group skipped significantly more words than the hard-question group of old adults, $F(1,53) = 4.2$, $MSe = 0.006$, $p < .05$.

In comparison to the easy-question group, the hard-question group of old adults showed prolonged single-fixation durations, $F(1, 53) = 5.7$, $MSe = 1115$, $p < .05$, second-fixation durations, $F(1, 53) = 8.2$, $MSe = 1004$, $p < .01$, gaze durations, $F(1, 53) = 10.7$, $MSe = 1897$, $p < .01$, as well as total reading times, $F(1, 53) = 32.2$, $MSe = 4466$, $p < .001$. Furthermore, a marginally significant difference between the two old groups was found in fixation position: Single fixations were located further to the left within words in the hard-question compared to the easy-question condition, $F(1, 53) = 4.0$, $MSe = 0.002$, $p = .05$.

In Table 2, means of word properties and duration of first-pass single fixation cases, that constitute the largest part of fixation types in all four groups (cf. Table 1), are listed for each group. Again, there was no significant difference in mean single fixation duration and no significant difference in selection effects between the two young groups (all ps > .05). In contrast, in line with the reduced skipping rate, comparison of the two groups of old adults revealed clear selection effects: Words that received a single fixation were shorter, $F(1,53) = 15.8$, $MSe = 0.002$, $p < .001$, higher in frequency, $F(1,53) = 15.9$,

$MSe = 0.03$, $p < .001$, and higher in predictability, $F(1,53) = 16.2$, $MSe = 0.01$, $p < .001$, in the hard-question than in the easy-question group of old readers. In line with these differences, the proportion of fixated function words was significantly larger for single-fixation cases of the hard-question than the easy-question group of adults, $F(1,53) = 15.8$, $MSe = 0.004$, $p < .001$. Interestingly, the mean single-fixation durations did not differ between easy-question and hard-question groups of old adults ($p > .05$; see Table 2).

(Table 2 about here)

Foveal and parafoveal word frequency effects

The final LMM, fitting log single fixation duration (SFD) on word n, was based on 21,738 fixations for young readers and 18,551 fixations for old readers. To test our hypothesis that selection effects influence the size of parafoveal-on-foveal effects, we tested cross-level interactions between reading condition and word frequency of the fixated word n and word frequency of the upcoming word n+1 in the two age groups. A list of estimates of variance components and fixed effects is provided in the Appendix. In the following, we focus on the results that pertain to our experimental hypotheses. All effects were significant with simultaneous statistical control of all the other effects listed in the Appendix.

In young readers, both groups showed the expected relationship of foveal word frequency and SFD as illustrated in Figure 1 (left panel). Generally, SFD decreased with increasing word frequency, resulting in significant linear ($b = -5.016$, $SE = 1.350$, $t = -3.71$), quadratic ($b = -2.805$, $SE = 1.136$, $t = -2.47$), as well as cubic trends ($b = -5.317$, SE = 0.6398, t = -8.31) of log word frequency. The quadratic coefficient of word frequency was significantly smaller in the hard-question than the easy-question group of young adults (b = 2.86, SE = 0.7413, t = 3.86), but none of the other interactions between word frequency and reading condition was significant for the two young groups.

Single fixation duration on word n was significantly modulated by the frequency of the parafoveal word (see Figure 1, right panel). Young adults in the easy-question and the hard-question conditions exhibited reduced SFDs with increasing frequency of word n+1 (b = -0.0214, SE = 0.0038, t = -5.56); there was no significant interaction with reading condition.

(Figure 1 about here)

For old adults, the linear and cubic trends of the frequency of word n were reliable in the easy-question condition (linear: b = -5.055, SE = 1.058, t = -4.78; quadratic: b = 1.469, SE = 0.097, t = 1.61; cubic: b = -1.861, SE = 0.5506, t = -3.38). Overall, SFD decreased with increasing word frequency, but there was a slight increase in SFD again for words of highest frequency. The hard-question group of old adults differed significantly from the easy-question old group in the linear and cubic trend of word frequency (linear: b = 2.626, SE = 0.6482, t = 4.05; quadratic: b = 0.5285, SE = 0.6577, t = 0.80; cubic: b = -1.512, SE = 0.6485, t = -2.33). As illustrated in the left panel in Figure 2, the readers in the hard-question group showed a steep negative slope on SFD for low frequency words, but increasing SFD for high-frequency words. Thus, the general trend of decreasing SFD with increasing word frequency was reduced in the hard-question group of old adults.

The parafoveal-on-foveal word frequency effect is illustrated in the right panel in Figure 2. Although the easy-question group of old adults showed already a significant negative effect of upcoming word frequency on SFD (b = -0.0169, SE = 0.004, t =-4.22), the parafoveal-on-foveal word frequency effect in the hard-question group of old adults was even significantly stronger (b = -0.0134, SE = 0.0035, t =-3.84). This was mainly due to long SFDs for low frequent words n+1.

(Figure 2 about here)

Discussion

The goal of our study was to investigate the influence of fixational selectivity on the measurement of lexical parafoveal-on-foveal effects during sentence reading using multivariate statistical analysis. A range of empirical evidence for lexical and sublexical parafoveal-on-foveal effects has been reported and the interaction of word properties of the fixated word (foveal load) as well the visibility of the parafoveal word within the perceptual span have been discussed as an explanation (Kennedy & Pynte, 2005; Kennedy, et al., 2002; Kliegl, et al., 2006). As far as global eye-movement statistics are concerned, old readers were reported to differ from younger readers in their skipping probability during sentence reading (Kliegl, et al., 2004; Rayner, et al., 2006). This difference implies a selection effect for words that are available for the analysis of single-fixation durations. To trace the consequence of the global text statistic (i.e., selection of fixation of short and highly frequent words) for local parafoveal-on-foveal effects, we experimentally manipulated the reading behaviour of both young and old readers that had a systematic effect on fixational selectivity, especially in old adults. Using identical sentence material for all groups, we changed the reading style from a superficial to a deliberate reading strategy by using difficult comprehension questions compared to easy questions. For old adults the manipulation led to a significant change in the properties of words that were available for multivariate analyses of single fixation durations.

Age differences between easy-question and hard-question groups in foveal and parafoveal frequency effects on single fixation durations are interpreted with reference to these selection effects. There are three main results. First, the difficulty of comprehension questions impacts on reading strategy in first-pass reading — leading to more refixations and, most notably, to a reduction of skipping probability in old adults. Second, the selection of fixated words is a critical parameter in the dynamics of reading, linking reader-level and word-level effects. Third, selection effects are critical for tests of distributed processing with multivariate analyses of eye-movement corpora. They may be the source for some of the differences between multivariate analyses of fixation durations encompassing a wide variety of words and univariate analyses of target words from experiments involving gaze-contingent display changes, usually containing only longer content words. In the following paragraphs, we discuss each of these results.

There was clear evidence for top-down influences of reading style on eye movement behaviour during sentence reading. The demanding comprehension questions led to differences in reading strategies expressed in more second-pass reading as well as in a reduced response accuracy in both age groups; old readers were more strongly affected by the question manipulation than young readers, as seen in an increase in refixations and longer first-pass gaze duration. Furthermore, old readers' fixational selectivity in first-pass reading was clearly affected by the experimental manipulation. Old readers in the hard-question condition showed a reduced skipping probability and in line with fewer skippings, words associated with single-fixation cases in first-pass reading were shorter and higher in frequency and predictability compared to words associated with single-fixation cases in the easy-question group of old adults. The increase in the percentage of fixated function words along with a decrease in skipping rate was an expected result, because short and high frequency words are usually skipped during reading, and function words are more often skipped than content words (e.g., Drieghe, Pollatsek, et al., 2008).

The impact of top-down influences such as depth of reading on local eye movement behaviour had recently been reported (Radach, Huestegge, & Reilly, 2008) and brings up the methodological issue how we control for the readers' intention and/ or attention in reading experiments (Radach & Kennedy, 2004). In reading studies that investigate processes in normal reading, i.e. during reading for comprehension, some kind of comprehension question is usually asked after a certain proportion of trials. This is the case independent of whether hypotheses are tested with analyses of specific target words or with multivariate analyses of all words. One concern with the results is that questions in the easy condition were so easy that readers may have only skimmed the sentences. This interpretation, however, is not born out by benchmark effects in the two conditions. In all four groups, word-level predictors showed the expected pattern of word frequency, word length, and word predictability effects. Thus, there was sufficient processing during reading the sentences. Moreover, if subjects in the easy-question condition had only skimmed the sentences, we would expect even shorter fixation durations and a reduced word frequency effect (e.g., Rayner & Fischer, 1996).

The properties of the fixated words of the old readers in the hard-question condition were more similar to the word characteristics of both young groups than to their age-matched control group reading in the easy-question condition. We interpret this result as the consequence of a more deliberate reading strategy translating into a reduction of skipping rate. Indeed, the higher proportion of skipping in older readers reported in earlier research (Kliegl, et al., 2004; Laubrock, et al., 2006; Rayner, et al., 2006) may be an inclination of lapses of attention ("mindless reading"). However, old adults' high-level of accuracy in the easy-question condition is not in agreement with this explanation. Alternatively, the default mode of old adults' reading has been characterized as higher risk-taking (Rayner et al., 2006) or an age-related lack of resilience in modulating fixation durations in response to processing opportunities in the perceptual span (Risse & Kliegl, 2011). Currently, all three explanations are speculations in need of experimental tests.

Second, the hypothesis that selection effects impact on the size of parafoveal-on-foveal word frequency effects was clearly supported. A reliable frequency effect of the parafoveal word on single fixation duration was found in all four experimental groups. Single-fixation duration is longer, when the parafoveal word is lower in frequency, a result that is in line with previous findings (Kennedy & Pynte, 2005; Kliegl, et al., 2006). Moreover and critically, in the hard-question condition old readers fixated words that on average were shorter, higher in frequency, and more predictable than those recorded for the easy-question group and showed a stronger impact of parafoveal word frequency on single fixation duration. We interpret this increase in the parafoveal-on-foveal frequency effect as an indicator for reliable evidence due to preprocessing of the upcoming word given a higher prevalence of fixated short and high frequent function words. An increased preprocessing effect is compatible with the assumption of distributed processing in the perceptual span, and, possibly, its modulation by the difficulty of the fixated word, as predicted by the foveal-load hypothesis (Henderson & Ferreira, 1990). The evidence for a dynamic modulation of the perceptual span in interaction with foveal load is indirect and the actual size of the effective span should be determined by reading experiments using eye-contingent display techniques (McConkie & Rayner, 1975). Further, our result of a stronger parafoveal-on-foveal effect in association with an increase of fixated function words is not an argument against serial processing accounts. Even though we favour the parallel processing account, a model-supported explanation of this effect clearly needs further research and we will not pursue this argument further at this point.

Importantly, this parafoveal-on-foveal effect, as measured as a cross-level interaction between frequency of word n+1 and reading condition in LMM, was reliable while controlling for unspecific individual differences (random effect of subject ID), random effects due to words and sentences, individual differences in reading style (first and second principal components), as well as various other word-level predictors (e.g., word length, word predictability, fixation location, saccade amplitude). Furthermore, it is difficult to account for this result as an artifact of mislocated fixations (Drieghe, Rayner, & Pollatsek, 2008; Rayner, et al., 2007) because we restricted our analyses to single fixation cases and even more importantly, readers in the hard-question group of old readers tended to locate single fixations further to the left within a word, i.e., farther from the word to the right of fixation, which is opposite to what would be expected under the hypothesis of an increase in the probability of mislocated fixations (see also Kennedy, 2008, for a discussion).

On the other hand, one could argue that in the hard-question old group part of the single fixation cases on function words are mislocated refixations of word n-1 (overshoots) and therefore reflect ongoing processing of the previous word rather than preprocessing of the next word n+1. And because function words are mostly followed by content words, these unintended, prolonged fixations on function words could produce a correlation with the frequency of the next word[3]. This alternative explanation of an unintended fixation on word n would suggest that fixation duration should even be modulated more strongly by frequency of word n-1. However, at least in our data, the spill-over effect of word frequency from word n-1 on word n is smaller for the hard-question old group than the easy-question old group which is not in agreement with this alternative explanation (cf. Table A2).

Differences in foveal word-frequency effects between the old groups lend further support for the interpretation of more preprocessing due to a higher prevalence of easy foveal words. The cross-level interaction shows that the cubic trend of word frequency on single fixation duration was more pronounced in the hard-question old group, that showed even larger increased single fixation durations for very high frequent words compared to the easy-question group. Again, this increase in processing time on very easy words n, that were more often function words in the hard- compared to the easy-question group, may reflect preprocessing of word n+1.

The present results provide further evidence for the assumption of distributed processing within the perceptual span that is in line with previous research (Kennedy & Pynte, 2005; Kliegl, 2007; Kliegl, et al., 2006). Results of a multivariate analysis using LMM including relevant reader-level predictors and word-level predictors of word n-1, word n, and word n+1 revealed a significant influence of parafoveal word properties on the processing of word n, as reflected in single-fixation duration on word n. When including all words in the analysis, concerns about problems of multicollinearity, ignoring skipping status of neighbouring words and syntactic class have been raised before (Rayner, et al., 2007). In our analysis and interpretation of the results, we addressed these critical issues by explicitly focusing on selection effects between experimental reading conditions. First, we demonstrated that the prevalence of fixating high frequency words goes hand in hand with a higher proportion of fixated function words, and second, we argue that exactly this selection patterns allows for stronger parafoveal-on-foveal frequency effects, as demonstrated for the old readers. Therefore, when skipping word n+1, longer single fixation durations on high frequent foveal words and before low frequent parafoveal words would be expected. Our results with respect to selection support the analysis provided by Kliegl (2007) who found skipping costs in single fixation duration on function words before skipping a content word.

Thus, finding positive or negative evidence for parafoveal processing does not only depend on the selection of reading material, for example by controlling for length and frequency of foveal words (Kennedy, et al., 2002), but these effects also depend on individual (e.g., age-related) or experimentally induced differences in reading strategy. Multivariate statistics reveal experimentally induced selection effects that may give rise to differences in the size of parafoveal-on-foveal effects. Conversely, effects reported from research with the paradigm of gaze-contingent display change of target words will also strongly depend on the properties of the words in the target region. Thus, given the usual lack of function words in these analyses, chances of detecting parafoveal-on-foveal effects may be substantially reduced. The conclusion is that evidence for parafoveal-on-foveal effects depends on properties of the words in the perceptual span as well as on individual and possibly experimentally induced differences in reading strategy.

In summary, the results demonstrate that differences in reading strategy --age-related or experimentally induced by task demands-- are predictive of eye movement patterns during reading of isolated sentences. Task demands change eye-movement measures at the level of the reader (e.g., skipping and regression probability, fixational selectivity of function words, average fixation duration). Cross-level interactions are consistent with the explanation that specification of a global parameter of eye-movement control in anticipation of the difficulty of questions, entails a reading strategy that regulates skipping probability and in the end gives rise to differences in parafoveal-on-foveal effects. Thus, on-line processes in reading are susceptible to the depth of reading and the reader's age, possibly tied in with age-related efficiency of executive control processes, and individual differences in reading style. In this context, linear mixed models allow us to test simultaneously a large number of subject-related, sentence-related, and word-related factors that influence eye movements in reading in a coherent data-analytic framework and unveil the dynamics of processing contingencies in the perceptual span.

Appendix A: Linear Mixed Model Results

(Table A1 about here)

(Table A2 about here)

Footnotes

1. Data from the easy-question groups of young and old readers were included in the analyses in Kliegl et al. (2006) and Kliegl (2007), labeled as group 4 and group 9 respectively.

2. The ten subject-level predictors in the PCA were mean incoming and outgoing saccade amplitude, mean skipping probability of the previous and next word, and mean frequency, mean length, and mean proportion of fixated content words of the previous and fixated words.

3. Erik Reichle suggested this explanation.

Acknowledgements

This research is part of Christiane Wotschack's dissertation (supported in part by a doctoral fellowship of the University of Potsdam) and was carried out in the context of a European Collaborative Research Project grant (PIs: Reinhold Kliegl, Alan Kennedy, Richard Shillcock, and Heinz Wimmer), with the German leg funded by Deutsche Forschungsgemeinschaft KL 955/7 and KL 955/14; PIs: Reinhold Kliegl and Ralf Engbert). We would like to thank Erik D. Reichle and two anonymous reviewers for their helpful comments on an earlier version of the manuscript. Data and R scripts are available at Potsdam Mind Research Repository (http://read.psych.uni-potsdam.de/pmr2).

References

Bates, D. M., & Maechler, M. (2009). Lme4: Linear mixed-effects models using S4 classes. R package version 0.9993765-32.

Buswell, G. T. (1922). *Fundamental reading habits: a study of their development.* Chicago: University of Chicago Press.

Drieghe, D., Desmet, T., & Brysbaert, M. (2007). How important are linguistic factors in word skipping during reading? *British Journal of Psychology, 98*(Pt 1), 157-171.

Drieghe, D., Pollatsek, A., Staub, A., & Rayner, K. (2008). The word grouping hypothesis and eye movements during reading. *Journal of Experimental Psychology: Learning, Memory & Cognition, 34*(6), 1552-1560.

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.

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

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*, 7192-7197.

Geyken, A. (2007). The DWDS Corpus: A Reference Corpus for the German Language of the 20th century. In C. Fellbaum (Ed.), *Idioms and Collocations: From Corpus to Electronic Lexical Resource*. Birmingham: Continuum Press.

Heister, J., Würzner, K. M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., et al. (2010). dlexDB - Eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau, 62*(1), 10-20.

Heller, D. (1985). Eye movements in reading. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements* (pp. 23 - 40). Birmingham: 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, & 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, 47*(2), 201-221.

Huey, E. B. (1908). *The psychology and pedagogy of reading*. New York: Macmillan.

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

Kennedy, A. (2000). Parafoveal processing in word recognition. *Quarterly Journal of Experimental Psychology A, 53*(2), 429-455.

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

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

Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology A, 55*(4), 1307-1337.

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

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.

Laubrock, J., Kliegl, R., & Engbert, R. (2006). SWIFT explorations of age differences in eye movements during reading. *Neuroscience and Biobehavioral Reviews, 30*(6), 872-884.

Lehrl, S. (1977). *Mehrfach-Wortschatz-Test Form B [mulitiple choice vocabulary test]*. Erlangen, Germany: Perimed.

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

Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2009). Parafoveal magnification: visual acuity does not modulate the perceptual span in reading. *Psychological Science, 20*(6), 721-728.

O'Regan, J. K., & Levy-Schoen, A. (1987). Eye movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance (Vol 12): The Psychology of reading.* Hillsdale, NJ: Erlbaum.

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

R Development Core Team. (2009). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. *Psychological Research, 72*(6), 675-688.

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for future research. *Journal of Cognitive Psychology, 16*(1-2), 3-26.

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

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

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

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

Rayner, K., & Pollatsek, A. (1987). Eye movements in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance* (Vol. 12, 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; discussion 530-527.

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging, 21*(3), 448-465.

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

Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading. *Psychology and Aging*, DOI: 10.1037/a0021616.

Schad, D., Nuthmann, A., & Engbert, R. (2010). Eye movements during reading of randomly shuffled text. *Vision Research, 50*, 2600-2616.

Tinker, M. A. (1958). Recent studies of eye movements in reading. *Psychological Bulletin, 55*(4), 215-231.

Wechsler, D. (1964). *Der Hamburg-Wechsler-Intelligenztest für Erwachsene (HAWIE)*. Berlin, Switzerland: Huber.

Wotschack, C. (2009). *Eye movements in reading strategies: How reading strategies modulate effects of distributed processing and oculomotor control*. Potsdam: Universitätsverlag.

Table 1

Word-based summary statistics broken down by age group and reading condition

Variable		Easy-question young	Hard-question young	Easy-question old	Hard-question old
Fixation probabilities					
Skipping	M	.16	.16	.25	.21
	SD	.07	.07	.09	.06
Single fixation	M	.68	.64	.59	.59
	SD	.06	.06	.08	.06
Double fixation	M	.10	.11	.08	.10
	SD	.04	.03	.04	.04
Three-plus fixation	M	.02	.02	.01	.02
	SD	.01	.01	.01	.02
Regression	M	.07	.18	.14	.25
	SD	.04	.08	.10	.17
Relative fixation position					
Single fixation	M	.53	.52	.51	.49
	SD	.04	.04	.03	.05
1st of multiple	M	.25	.25	.40	.39
	SD	.06	.07	.15	.12
2nd of multiple	M	.66	.64	.53	.54
	SD	.06	.07	.12	.11
Fixation duration (ms)					
Single fixation	M	231	242	224	245

	SD	31	36	31	37
1st of multiple	M	211	216	218	233
	SD	23	29	29	37
2nd of multiple	M	190	203	184	208
	SD	26	29	32	31
gaze duration	M	261	277	250	289
	SD	40	42	37	51
Total reading time	M	281	358	279	383
	SD	47	76	52	84

Note. Data are from right eye and include all detected fixations. Probabilities of skipping, single, double, and three or more fixations are based on first-pass fixations preceded and followed by a forward saccade; thus, first-pass fixations that were regression origins were excluded.

Table 2

Means and standard deviations of word properties and fixation duration of first-pass single fixation cases broken down by age group and reading condition

Variable word n		Easy-question young	Hard-question young	Easy-question old	Hard-question old
Frequency (log/ million)	M	2.24	2.26	2.03	2.23
	SD	0.20	0.16	0.22	0.13
(PSC: M = 2.3, SD = 1.3)					
Predictability (logit)	M	-1.58	-1.55	-1.65	-1.52
	SD	0.12	0.10	0.12	0.10
(PSC: M = -1.48, SD = 1.1)					
Length (n of letters)	M	4.5	4.5	4.9	4.6
	SD	0.4	0.3	0.4	0.3
(PSC: M = 5.4, SD = 2.6)					
Function word proportion	M	0.32	0.33	0.27	0.33
	SD	0.06	0.04	0.07	0.04
(PSC: M = 0.37, SD = 0.48)					
Single fixation duration (ms)	M	220	226	220	237
	SD	24	30	32	42

Note. Data are from right eye. First and last fixations in a sentence and fixations on the first and last word were excluded from the analyses.

Table A1

Final linear mixed model fitting log single fixation duration for the easy-question and hard-question young groups, fit by maximum likelihood: Variances and standard deviations of random effects; means, standard errors, and t-values of fixed effects.

Random effects		
Groups	Variance	SD
Word ID	0.0099	0.0994
Sentence ID	0.0019	0.0435
Subject ID	0.0125	0.1117
Residual	0.0850	0.2915
Number of obs.: 21738, groups: words 550, sentences144, subjects 54		

Fixed effects			
	Estimate	SE	t-Value
Intercept	5.3410	0.0283	188.89
Trial	-0.0004	0.0001	-5.81
Vocabulary	-0.0127	0.0059	-2.16
Condition (cnd)	0.0017	0.0332	0.05
PC1	0.0174	0.0058	2.99
PC2	-0.0311	0.0124	-2.51
Poly(frequency(n))1	-5.0160	1.3500	-3.71
Poly(frequency(n))2	-2.8050	1.1360	-2.47
Poly(frequency(n))3	-5.3170	0.6398	-8.31
Predictability(n)	-0.0198	0.0036	-5.55
1/length(n)	0.0236	0.0870	0.27
Length 6,7 letters	-0.0498	0.0115	-4.32

Frequency(n-1)	-0.0366	0.0039	-9.43
Predictability (n-1)	-0.0138	0.0038	-3.63
1/length(n-1)	0.3296	0.0430	7.67
Frequency(n+1)	-0.0214	0.0038	-5.56
Predictability(n+1)	0.0117	0.0035	3.30
1/length(n+1)	0.1030	0.0414	2.49
Incoming sacc. ampl.	0.0329	0.0014	23.83
Rel. fix. position	-0.1097	0.0154	-7.12
(Rel. fix. position)2	-0.3204	0.0369	-8.69
Outgoing sacc. ampl.	0.0132	0.0016	8.49
Length*freq.(n)	0.1963	0.0638	3.07
Freq.(n-1)*freq.(n)	0.0128	0.0021	6.27
Freq.(n)*freq.(n+1)	0.0077	0.0022	3.58
Cnd*trial	0.0003	0.0001	3.00
Cnd*poly(freq(n))1	-1.2780	0.7796	-1.64
Cnd*poly(freq(n))2	2.8600	0.7413	3.86
Cnd*poly(freq(n))3	-0.1489	0.6102	-0.24
Cnd*1/length(n)	0.1415	0.0640	2.21
Cnd*predictability(n+1)	0.0084	0.0034	2.52
Cnd*incoming sacc. ampl.	-0.0070	0.0018	-3.99
Cnd*rel. fix. position	0.0742	0.0188	3.95
Cnd*outgoing sacc. ampl.	0.0133	0.0022	6.10
Cnd*length(n)*freq(n)	-0.1642	0.0468	-3.51

Table A2

Final linear mixed model fitting log single fixation duration for the easy-question and hard-question groups of old adults, fit by maximum likelihood: Variances and standard deviations of random effects; means, standard errors, and t-values of fixed effects.

Random effects		
Groups	Variance	SD
Word ID	0.0069	0.0830
Sentence ID	0.0019	0.0437
Subject ID	0.0185	0.1361
Residual	0.0832	0.2885
Number of obs.: 18551, groups: words 550, sentences 144, subjects 55		

Fixed effects			
	Estimate	SE	t-Value
Intercept	5.4070	0.0344	157.23
Trial	-0.0003	0.0001	-4.55
Vocabulary	-0.0256	0.0146	-1.75
Condition (cnd)	0.0023	0.0428	0.05
PC1	0.0012	0.0078	0.15
PC2	-0.0333	0.0130	-2.56
Poly(frequency(n))1	-5.0550	1.0580	-4.78
Poly(frequency(n))2	1.4690	0.9097	1.61
Poly(frequency(n))3	-1.8610	0.5506	-3.38
Predictability(n)	-0.0201	0.0036	-5.54
1/length(n)	0.1122	0.0709	1.58
Length 6,7 letters	-0.0432	0.0103	-4.18

Frequency(n-1)	-0.0354	0.0042	-8.52
Predictability(n-1)	-0.0073	0.0039	-1.89
1/length(n-1)	0.2466	0.0442	5.57
Frequency(n+1)	-0.0169	0.0040	-4.22
Predictability(n+1)	0.0066	0.0031	2.11
1/length(n+1)	0.1766	0.0418	4.23
Incoming sacc. ampl.	0.0272	0.0010	27.4
Rel. fix. position	-0.0267	0.0119	-2.25
(Rel. fix. position)2	-0.1679	0.0364	-4.61
Outgoing sacc. ampl.	0.0089	0.0012	7.26
Freq(n-1)*freq(n)	0.0093	0.0020	4.59
1/length(n)*freq(n)	0.1584	0.0539	2.94
Freq(n)*freq(n+1)	0.0059	0.0021	2.78
Cnd*trial	0.0004	0.0001	3.54
Cnd*poly((freq(n))1	2.6260	0.6482	4.05
Cnd*poly((freq(n))2	0.5285	0.6577	0.80
Cnd*poly((freq(n))3	-1.5120	0.6485	-2.33
Cnd*freq(n-1)	0.0084	0.0036	2.35
Cnd*freq(n+1)	-0.0134	0.0035	-3.84
Cnd*incoming sacc. ampl.	-0.0060	0.0016	-3.77
Cnd*outgoing sacc. ampl.	0.0079	0.0020	3.90

Figure 1: Effects of frequency of word n (left panel) and word n+1 (right panel) on single fixation duration for the easy-question (solid line) and hard-question (dashed line) groups of young adults. Error bars represent within-subject 95% confidence intervals.

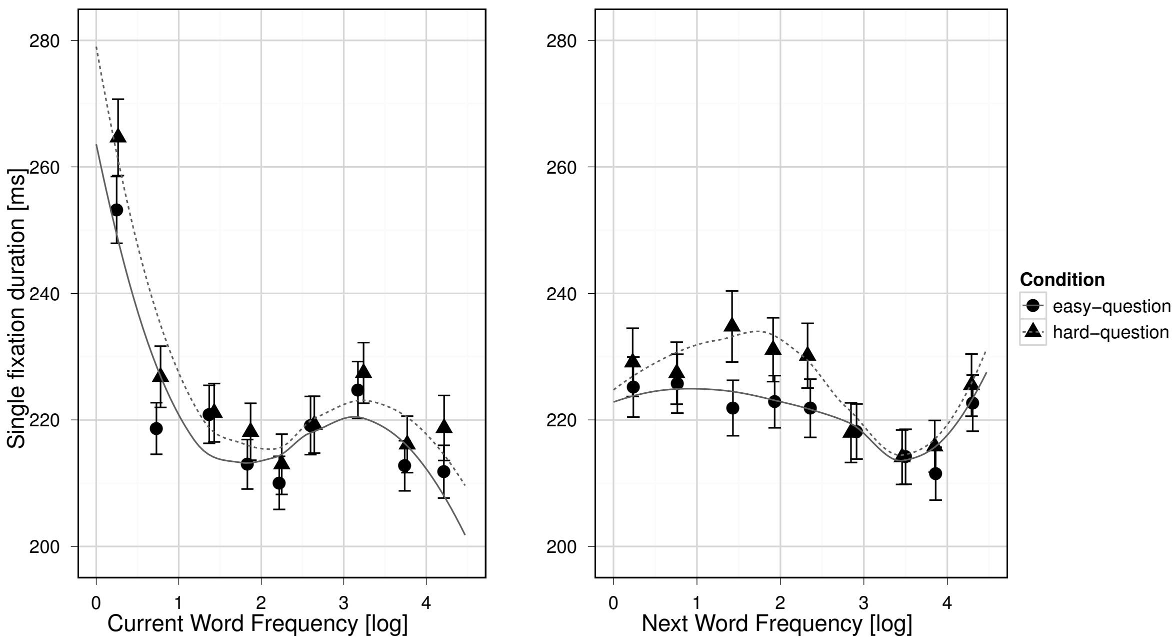

Figure 2: Effects of frequency of word n (left panel) and word n+1 (right panel) on single fixation duration for the easy-question (solid line) and hard-question (dashed line) groups of old adults. Error bars represent within-subject 95% confidence intervals.