The effect of word position on eye-movements

in sentence and paragraph reading

Victor Kuperman*

Stanford University, USA

Michael Dambacher

University of Potsdam, Germany

Antje Nuthmann

University of Edinburgh

Reinhold Kliegl

University of Potsdam, Germany

December 7, 2009

Running Head: Word position effect

*Corresponding author: Victor Kuperman, Stanford University, Margaret Jacks Hall, Building 460, Stanford CA 94305-2150, USA. Email: vickup@stanford.edu. Phone: 1-650-723-4284. Fax: 1-650-723-5666

Abstract

The present study explores the role of the word position-in-text in sentence and paragraph reading. Three eye-movement data sets based on the reading of Dutch and German unrelated sentences reveal a sizeable, replicable increase in reading times over several words in the beginning and the end of sentences. The data from the paragraph-based English-language Dundee corpus replicate the pattern and also indicate that the increase in inspection times is driven by the visual boundaries of the text organized in lines, rather than by syntactic sentence boundaries. We argue that this effect is independent of several established lexical, contextual and oculomotor predictors of eye-movement behavior. We also provide evidence that the effect of word position-in-text has two independent components: a start-up effect arguably caused by a strategic oculomotor program of saccade planning over the line of text, and a wrap-up effect originating in cognitive processes of comprehension and semantic integration.

Keywords: eye movements; word processing, sentence processing

ACKNOWLEDGMENTS

INTRODUCTION

It is a well-established finding in the literature on eye movements in reading that the time spent by the eyes on a word is indicative of the difficulty of processing that word. Experimental and corpus-analytical research has established a number of *benchmark* factors that robustly emerge as strong co-determiners of eye-movement measures, including word length, word frequency, and word predictability given preceding context (for an overview of literature, see Rayner, 1998): e.g., shorter, more frequent and more predictable words are read faster and are skipped more frequently. The present paper **aims to extend** the benchmark list by introducing word position-in-text as a reliable and strong predictor of reading times. To our knowledge, this is the first systematic cross-corpora and cross-linguistic study of word position effect on normal (mindful) and mindless reading of sentences.

Current research reports mixed findings on whether and how word processing is affected by the position of that word on a screen, in a sentence, or in a line of text. The controversy is evident in studies of reading both at the level of isolated sentences and at the level of paragraphs. For instance, several experiments report readers' tendency to speed up as they proceed through a sentence (Aaronson & Scarborough, 1976; Aaronson & Ferres, 1983; Chang, 1980; Ferreira & Henderson, 1993). These results, obtained using a variety of experimental techniques, including self-paced reading and eye-tracking, are argued to reflect the special role of initial words in a sentence (or initial sentences in a passage) as foundations for creating the mental representation of a larger unit (Gernsbacher, 1990). Since laying such a foundation is a cognitively demanding task, sentence-initial words are arguably processed

slower.

These results contrast with a robust finding of the wrap-up effect in sentence reading: Clause- or sentence-final words are read slower than the identical or matched words within a clause (cf. Balogh, Zurif, Prather *et al.*, 1998; Rayner, Sereno, Morris *et al.*, 1989; Rayner, Kambe & Duffy, 2000). Three reasons have been proposed for the clause-final inflation of reading times. One is that clauses represent processing units of sentence comprehension, and the end of the clause is the point where semantic integration of words within the clause and integration with prior discourse take place (Just & Carpenter, 1980). The other reason is that punctuation marks (commas, periods or longer spacing occurring at the end of clauses and sentences) may give rise to a low-level hesitation response of the oculomotor system (Hill & Murray, 2000; Warren, White & Reichle, 2009). The third proposed reason for the wrap-up effect is that in silent reading clause endings are marked by a specific contour of implicit prosody (Hirotani, Frazier & Rayner, 2006). Experimental data and simulations of the E-Z Reader model of sentence reading (Reichle, Warren & McConnell, 2009; Warren et al., 2009) appear to show that some or all of these reasons conspire in causing the inflated reading times at the clause ending. Importantly, all studies of the wrap-up effect concentrate on the processing of the last one or two words in the clause, while making no claim as to whether word position elicits any effect in the beginning or the middle of the sentence.

In the first part of the present paper we investigate whether moving towards the end of the string of text elicits a speed-up or a slow-down in processing, and what the loci of the effects are. We set out to explore the effect of word POSition in Text (henceforth, POST) in three

eye-movement corpora based on reading of unrelated sentences in German and Dutch, and on a paragraph-based English-language corpus. We will demonstrate that the word's horizontal position in a text has a robust and sizeable effect on sentence and paragraph reading, **such that inspection times are relatively short in the beginning of a single-line sentence or a line of text on the multi-line screen, and the reading times increase as the eyes move further into the sentence or line. Equivalently, the slope of the change in reading times increases with the rightward progression into the sentence or line.** We will also argue that the effect has two independently motivated components, the gradually increasing reading times as the eye departs from the beginning of the line and moves downstream (*start-up*), as well as the increase in reading times towards the end of the line (*wrap-up*)[1].

The second part of the paper adresses a range of potential mechanisms underlying the word position-related effects. First, it is possible that the POST effect might in fact reflect the influence of some lexical confound(s) known to affect processing times. For instance, word position in a sentence often serves as an index of the contextual constraint that the processed part of the sentence imposes on upcoming words. Generally, the further into the sentence a word is, the more predictable it is from its context. This intuition is corroborated by findings of several experiments, including those measuring eye-movements and event-related potentials (ERPs). For example, Van Petten and Kutas (1990) observed smaller amplitudes on the N400 component with increasing word position, which they interpreted as an indication of higher predictability and easier semantic integration of words at the end of

6

the sentence (cf. Van Petten & Kutas, 1991). Similarly, Dambacher, Kliegl, Hofmann and Jacobs (2006) used the Potsdam Sentence Corpus of 144 sentences with predictability norms collected for each word (cf. Kliegl, Grabner, Rolfs & Engbert, 2004) to establish independent contributions of word position and word's contextual predictability to the amplitudes of P200 and N400 components. Dambacher *et al.* (2006) attested significant positive correlations between word position in sentence, word predictability and word frequency in the Potsdam Sentence Corpus. Moreover, Dambacher et al. (2006) observed a significant correlation of word position-in-sentence with N400 amplitudes: the amplitudes decreased the further the readers were into the sentence. Yet the effect of word position did not reach significance in the statistical model that also included word frequency and the interaction of predictability by frequency, which implies that the two benchmark factors of word frequency and predictability subsumed all variance explained by word position. In light of such findings, we test here whether the POST effect is independent of word predictability as well as multiple other lexical covariates (word frequency, length, and lexical status of the currently fixated as well as of the preceding and the subsequent words), by statistically disentangling their respective contributions to predicting inspection times.

Second, the low-level oculomotor behaviour that readers demonstrate while reading strings of characters on the screen might emerge as a non-lexical cause for the POST effect. We will complement our corpus analytical approach with evidence obtained from an experimental manipulation, the so-called mindless reading paradigm (Nuthmann & Engbert, 2009; Rayner & Fischer, 1996; Vitu, O'Regan, Inhoff & Topolski, 1995). In this paradigm,

participants read sentences in both their normal version as well as a transformed (or mindless) version where each letter is replaced with a *z*. Such experimentation approximates reading in the sense of the visual processing of punctuated strings of symbols, yet it is devoid of word recognition or higher-level language processing. Thus, comparing data obtained under normal reading conditions vs. reading of z-strings will allow us to assess whether the POST effect is solely driven by non-lexical low-level visuomotor variables.

Third, word position in sentence, in line, and on the screen are naturally confounded in sentence reading. Hence eye-movement records for larger texts, where sentences and lines are commonly misaligned, might shed light on whether it is the visual object (line or screen) or the syntactic object (sentence or a paragraph) that is primarily responsible for the POST effect. Yet the current literature on paragraph reading shows contradictory findings. A recent study of the English and French parts of the paragraph-based Dundee corpus (Pynte and Kennedy, 2006) **reported that later positions in a line were associated with increased saccade amplitudes** (for French), a decrease in the skipping rate (for English and French), and a non-linear increase in inspection times (first fixation duration for English and French, and gaze duration for French). Conversely, Demberg and Keller (2008) observed a strong *negative* correlation of word position-in-sentence with the time it took to read that word in the English part of the Dundee corpus (first fixation, first pass duration and total reading time). The findings of two further studies on the French part of the corpus are mixed too: the increasing position-in-sentence elicited a significant increase in gaze duration for content words in Pynte, New and Kennedy (2008a), while eliciting a weak,

statistically non-significant decrease in gaze duration in Pynte, New and Kennedy (2008b). However, word position-in-line enters into a significant interaction with the syntactic indices of the contextual constraint in both Pynte, New and Kennedy (2008a) and Pynte, New and Kennedy (2008b).

To sum up, there is no consistent pattern of results coming from paragraph-reading experiments with respect to either word position-in-sentence, or its position-in-line. It is at present unclear whether the discrepancies in those studies stem from the differences in languages under study, data collecting techniques, statistical analytical techniques, the range of predictors used in the statistical models, or the focus on the sentence vs. the line as a unit of analysis. Below we present results that may shed light on the mixed reports in the current literature. We consider the English part of the Dundee corpus to disentangle the processing of visual objects (position-in-line) from that of syntactic objects (position-in-sentence) as predictors of inspection times. Finally, we propose a mechanism that may account for the observed behavioral patterns by relying on a visuo-oculomotor program of processing a line of text, as well as a higher-level semantic process of semantic integration, with both types of processes being triggered by visual characteristics of the text under inspection.

METHOD

Single-line sentence reading data

We set out to consider eye-movement data available for three corpora, each consisting of unrelated single-line sentences: the German-language Potsdam Sentence Corpus (henceforth, PSC), the Dutch Eye-Movements ONline Internet Corpus (henceforth, DEMONIC) and a

large-scale experiment on reading of Dutch sentences with embedded morphologically complex words (henceforth, DMORPH); see Table 1 for the description of the corpora, Kliegl, Nuthmann and Engbert (2006) for detailed specifications of PSC and Kuperman, Bertram and Baayen (in press) for detailed specifications of DMORPH. Sentences in all corpora were declarative and in most cases syntactically simple. The readers' task was reading for comprehension and answering multiple choice comprehension questions appearing after about 27% of sentences in PSC and yes-no questions after about 17% of sentences in DEMONIC and DMORPH. The experimental manipulations implemented in target words differed across corpora.

INSERT TABLE 1 APPROXIMATELY HERE

Table 1: Description of Eye-Movements Corpora

Language	Corpus	Sentences (words)	Readers	Sentence length in words	Single-fixation duration	Gaze duration
German	PSC	144 (1138)	222	5-11, $\mu = 8.2$ (1.4)	$\mu = 210$ ms (67)	$\mu = 233$ ms (106)
Dutch	DEMONIC	224 (1746)	55	5-20, $\mu = 10.9$ (2.7)	$\mu = 241$ ms (85)	$\mu = 268$ ms (118)
Dutch	DMORPH	158 (1421)	28	6-17, $\mu = 11.6$ (2.2)	$\mu = 237$ ms (82)	$\mu = 262$ ms (93)

Columns 5-7 report in parentheses standard deviations for corresponding mean values μ.

Recording procedures and the apparatus differed slightly across the two labs in which data were collected for PSC, and the lab in which data were collected for DEMONIC and DMORPH. The experimental setup for PSC used a 21-inch monitor size and a chin rest, while a 17-inch monitor and no chin rest were used in DEMONIC and DMORPH. Recordings and nine-point grid calibrations in PSC were done binocularly, while in DEMONIC and DMORPH recordings and three-point horizontal grid calibrations were based on the movements of the right eye. EyeLinkI and EyeLinkII were used for data collection in PSC (see

Kliegl, Nuthmann & Engbert, 2006) with the 250 Hz and 500 Hz sampling rates respectively, while in DEMONIC and DMORPH we recorded eye-movements using EyeLinkII with the 500 Hz sampling rate. Across corpora, sentences were presented one-by-one on the screen in font regular Courier New, and each sentence occupied exactly one line (80 or less characters). Sentences were presented in PSC such that the fixation mark was between the beginning and the middle of the first word, while in DEMONIC and DMORPH the fixation mark was placed 20 pixels to the right of the left screen edge and coincided with the initial character of the first word of the sentence. Readers were positioned approximately 80 cm from the monitor in DEMONIC and DMORPH and approximately 60 cm in PSC, such that one character subtended approximately $0.36°$ of the visual angle in DEMONIC and DMORPH, while one character subtended about $0.35°$ of the visual angle in PSC.

Independent Cloze sentence completion studies have been carried out to collect predictability norms for each word in PSC (83 complete predictability protocols) and, separately, for DEMONIC (50 complete predictability protocols). The norming study was administered in the laboratory setting for PSC, and as a web-based experiment for DEMONIC. In both studies, participants were instructed to guess the first word of the unseen original sentence and to enter it via the keyboard. The computer responded with displaying the first word of the original sentence on the screen. Then participants entered their guess for the second word, and so on, until the period appeared indicating the end of the sentence. Words from the original sentence stayed on the screen. Predictability was measured as the probability of correctly predicting a word after having seen the preceding part of the sentence (for details

of the norming procedure see Kliegl et al., 2004).

Data processing based on fixation sequences was identical across corpora. We excluded sentences with track losses and blinks, first and last fixations in sentences, fixations on first or last words in sentences, fixations shorter than 30 ms or longer than 1,000 ms, and fixations preceded or followed by microsaccades (i.e., within-letter saccades). In all corpora, the trimming procedure led to the loss of approximately 30% of data points (20% due to exclusion of sentence-extremal fixations, and 10% due to other selection criteria). In this study we only considered fixations made during the first-pass reading: **that is, we excluded regressions and second-pass fixations from each sentence, while other fixations in the sentence remained part of the analysis.** The remaining data pools consisted of 70,679 data points for PSC, 58,854 for DEMONIC and 22,769 for DMORPH.

Paragraph reading data

To explore whether patterns typical of sentence reading generalise over paragraph reading, we considered the English component of the Dundee corpus (Kennedy & Pynte, 2005). The Dundee corpus consists of a series of 20 newspaper texts, each comprising 2800 words, silently read by 10 participants. The texts were split for presentation into 40 five-line screens, while the maximum line length was 80 characters. The end of any line on the screen did not necessarily represent the end of the sentence, nor did the sentence-initial words coincide with line-initial words. **Also, the end of the screen did not necessarily coincide with the end of the paragraph.** The apparatus used for data collection was a Dr. Boise eye-tracker, which recorded the movements of the right eye with a sampling rate of 1 ms and

a spatial accuracy of 0.25 characters (for the full description of data collection method and stimuli, see Kennedy & Pynte, 2005).

The corpus included a large number of very long sentences (maximum = 87 words; median = 26 words). To ensure comparability of the analyses for the Dundee corpus data and our sentence-based corpora, we confined our data pool to the 5-15 word-long sentences. Also, we excluded fixations that landed on the first or the last word of a line or of a sentence for compatibility with other data sets and to avoid the potential influence of the eye-movement behaviour at line breaks. We additionally removed fixations that were shorter than 50 ms or longer than 1000 ms. Finally, we restricted our data to the first reading pass. These trimming procedures left us with 25,350 data points.

Reading of z-strings

Further, we report data from a sentence reading experiment that was created using the Potsdam Sentence Corpus (see PSC) (Nuthmann, Engbert & Kliegl, 2007; Nuthmann & Engbert, 2009). For a given sentence, all words were replaced with their corresponding z-strings, while the upper/lower case of characters and punctuation marks were preserved. For example, the z-string equivalent to the sentence `Reading is a complex skill.` would be `Zzzzzzz zz z zzzzzzz zzzzz.`. Participants were instructed to scan the z-strings as if they were reading. Z-string scanning shares the visuo-oculomotor requirements with reading, but none of the higher-level language-related processes. Thus, the paradigm can provide evidence about the role of lower-level visuo-oculomotor processes in the absence of higher-level cognitive operations generally involved in word recognition and sentence integration. Forty-

six participants contributed to the z-string scanning condition, while 26 of these participants also read the corpus sentences of PSC (for a full description of the experimental design and setup see Nuthmann et al., 2007; Nuthmann & Engbert, 2009).

Statistical considerations

For statistical analyses, we fitted experimental data with linear mixed models with participant, word, and sentence as random effects and a large set of fixed effects to test the independence of the POST effect against the backdrop of many control factors; for the full list of predictors and their effects see Table 2. All our models were trimmed such that the individual points that fell outside the range of -2.5 and 2.5 SD of the residual error of the model were excluded from consideration (below 2% of data in any of the models), and the models were re-fitted to the remaining subsets. Table 2 reports estimates of the regression coefficients, their standard errors and t-values. For the effects reported in the body of the paper we also provide beta coefficients and p-values estimated by the Monte Carlo Markov chain (MCMC) method using 10,000 samples (for a detailed treatment of the method, see Baayen, 2008; Baayen, Davidson & Bates, 2008; Pinheiro & Bates, 2000).

RESULTS AND DISCUSSION

The word position-in-sentence effect

The absolute word position (i.e., the word rank in the sentence) had a strong positive effect on both single-fixation duration and gaze duration in all three corpora; see left panels of Fig. 1 for single-fixation duration data. Words at the end of the sentence were fixated some 30-50 ms longer than those in the beginning of the sentence. We also zoom in on the

sentence-final increase in reading times by plotting the absolute word position aligned at the last word of the sentence (right panels of Fig. 1).

INSERT FIGURE 1 APPROXIMATELY HERE

Since sentences are of different lengths, most positions in the middle and the end of the horizontal range represent a mixture of sentence-medial and sentence-final words. For instance, the mean reading time for position 5 reflects both fixation durations for sentence-final words in sentences with 6 or 7 words, as well as fixation durations for sentence-medial words in sentences consisting of 10 to 12 words. The absolute word position in sentence does not allow for disentangling these potentially different strata of data. Therefore, the effect of word position was assessed separately for each sentence length (corresponding plots not shown in consideration of space). Single fixation and gaze duration increased most substantially in two regions: the words in the beginning of the sentence (a 10-20 ms increase across corpora) and words at the end of very long sentences (a 20-40 ms increase across corpora). Moreover, those sentence-initial and sentence-final slow-downs were present throughout the range of sentence lengths, and were not confined to just short or just long sentences.

As another way of controlling for variability in sentence length, we considered the word's *relative* position-in-sentence, defined as the ordinal rank divided by the sentence length in words and ranging from 0 to 1. Relative word position also takes into account the negative correlation of sentence length with single fixation duration, which reached statistical significance in DEMONIC and DMORPH (ps < 0.01) and was reported earlier by Aaronson and Scarborough (1976). It also confirms the intuition that word 3 in a sentence with only
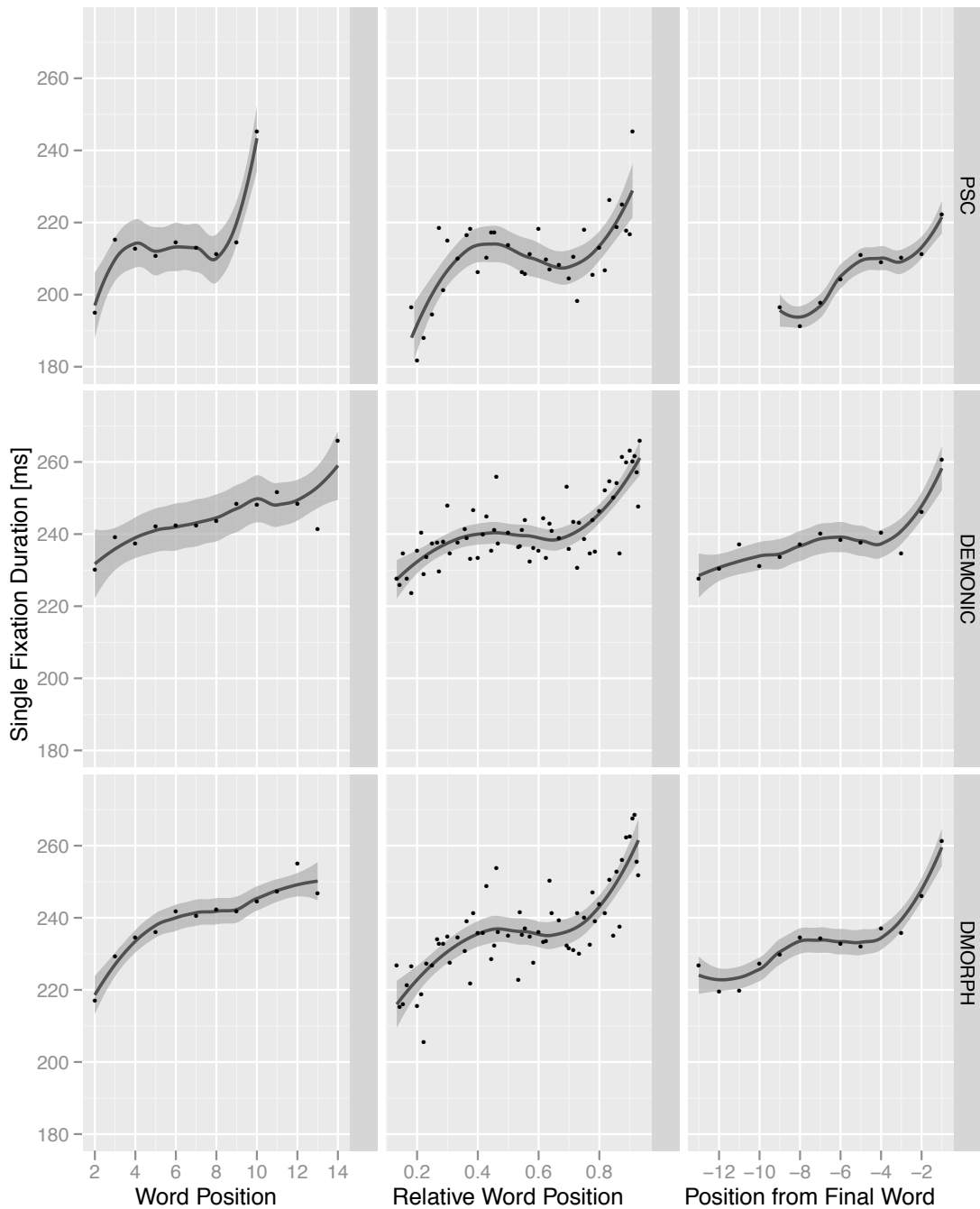
15

Figure 1: Single Fixation Duration as a function of absolute word position in sentence (left panels), relative word position in sentence (central panels) and absolute word position aligned at the sentence-final word, for data sets PSC, DEMONIC and DMORPH. Dots represent mean values of single fixation duration. The gray area marks the 95% confidence interval. The solid lines are approximations of the data by the polynomial spline with 3 knots.

5 words (relative position $\frac{3}{5} = 0.6$) may differ in the relative ease of semantic integration into the sentence meaning than, say, word 3 in a 15 word-long sentence (relative position $\frac{3}{15}$ = 0.2): the longer the sentence the more costly it may be for the reader to keep the beginning of the sentence in the short-term memory and the more material there is to integrate (see Gibson, 1998 for a similar reasoning about linguistic complexity, **see also Patson & Warren (in press) for evidence against locality effects in semantic and discourse processing**). Accordingly, the central panels of Figure 1 plot the effect of relative word position-in-sentence on single fixation duration across three eye-movement data sets. Plots for gaze duration as a function of the absolute and relative word position-in-sentence (not shown) reveal very similar patterns. As an aside, we note that the relative POST curve appears to be more similar across the three sets of data than the absolute POST curves, for both single fixation duration and gaze duration.

Again, across the three corpora, relative word position-in-sentence elicited an inflation of single fixation and gaze durations, with a 30-45 ms average difference between the second and the penultimate words of sentences. Plots of the relative word position support the notion that extremes of the sentence are the regions where the inspection times increase faster, even as the first and last words of sentences were excluded from our consideration. This, to our knowledge, is the first time that the sentence-initial increase in inspection times has been described. We also report a novel finding that the inflation in reading times is not confined to the last word in a clause or a sentence investigated in early studies, but rather increases in magnitudes over several words at the end of the sentence. We dub the sentence-initial

increase in reading times the 'start-up' effect, and use the established label of the 'wrap-up' effect for the sentence-final increase (see above for our adoption of the term 'wrap-up').

To sum up our results so far, the word's POST shows a substantial and robust effect replicable across different languages (Dutch and German), different sets of experimental stimuli and different populations of participants. The functional relation between (absolute or relative) word position in sentence and fixation durations for words has a shape of the cubic parabola, as suggested by the interpolating spline function plotted as a solid curved line in Figure 1. It is a logical possibility that inspection times enter into a complex functional (e.g., cubic polynomial) relationship with word position, or some processing factor that word position is an index of. The alternative that we argue for below is that this parabolic shape is a juxtaposition of the increases in reading times at the extremes of word sequences, and that these two behavioural patterns are likely the outcomes of different processing phenomena. The next section complements the data patterns observed in single-line sentence reading by the analysis of the paragraph-based Dundee corpus.

Visual processing of paragraphs

In paragraph reading, word position-in-sentence and position-in-line are generally misaligned, so we checked both possibilities for replication of the increase in inspection times. Plots of absolute or relative word position-in-sentence against inspection times in the Dundee corpus did not reveal any obvious patterns, even for the subset of the data which comprised 5-15 word-long sentences; see Figure 2 for the scatterplot of single fixation duration against the relative word position-in-sentence. The mean single fixation showed only a negligible

18

variability (within 10 ms) throughout the sentence (except for sentence length 6), as indicated by the lowess smoother lines in Figure 2. Likewise, word position-in-sentence did not elicit a significant effect in the linear mixed model for either single fixation duration or gaze duration (all $ps > 0.1$; the model also included frequency and length of words N-1, N and N+1, and the line on screen as fixed effects, as well as word, participant, screen and sentence as random effects). This finding is at odds with the report of Demberg and Keller (2008) that the increasing position-in-sentence comes with faster inspection times: most likely, the effect they observed was driven by extremely long sentences.

INSERT FIGURE 2 APPROXIMATELY HERE

Figure 3 summarizes the effect of relative word position-in-line on single fixation duration in the Dundee corpus, broken down by the line number (1 to 5) on the five-line screen. Plots with absolute word position-in-line on the x-axis and with gaze duration on the y-axis revealed very similar qualitative patterns (not shown here in consideration of space).

INSERT FIGURE 3 APPROXIMATELY HERE

Figure 3 reveals the familiar functional shape of the cubic parabola – with the steeper line-initial and line-final inflation of single fixation duration – at the last, fifth, line on the screen, such that the penultimate word on the 5th line is about 40 ms longer than the second word on that line. That is, last lines on the screen are processed in a manner qualitatively similar to the reading of single-line sentences. The start-up effect is actually present in all lines (1-5), since words in the line-initial region were read somewhat faster than those in the
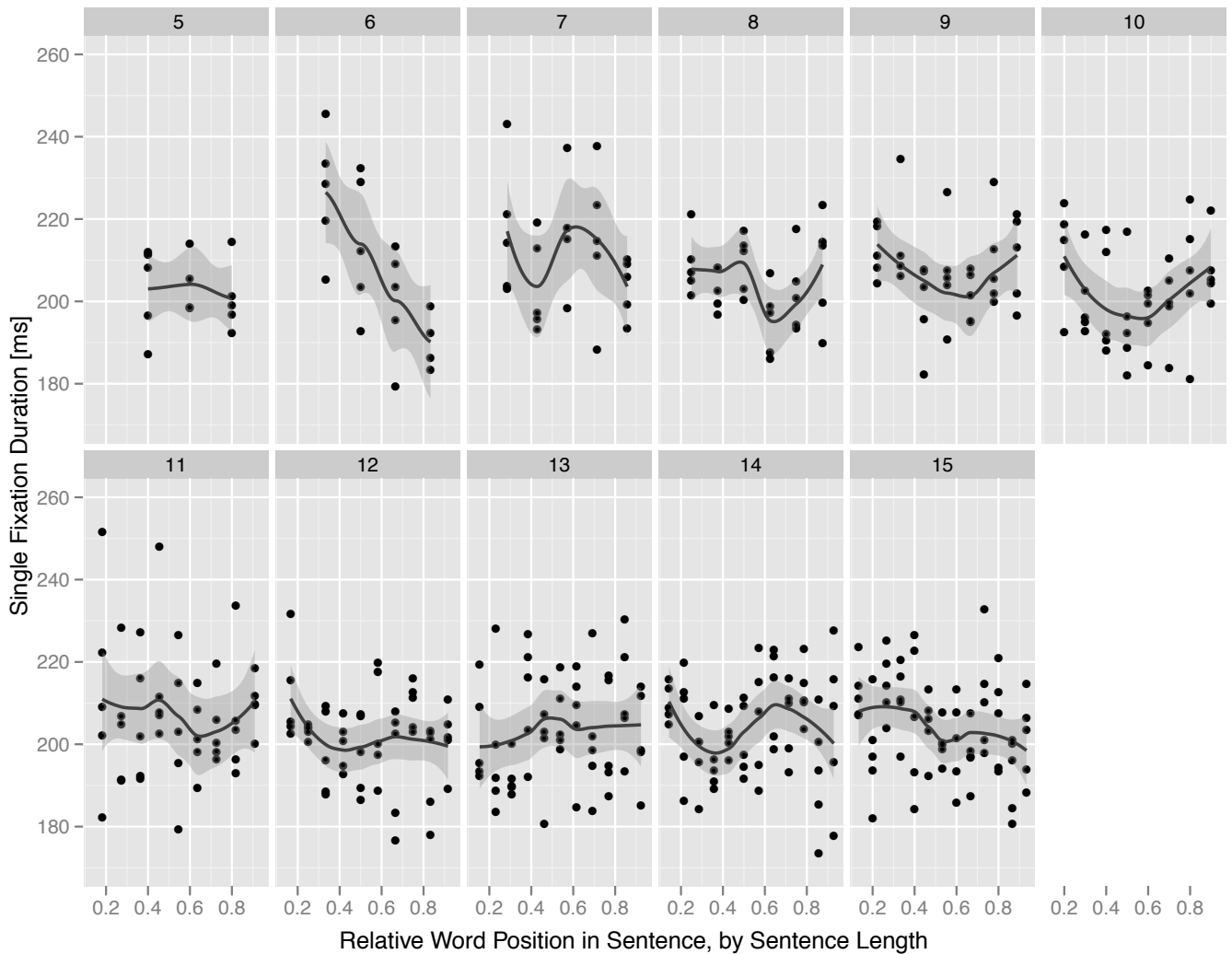
19

Figure 2: Single Fixation Duration as a function of relative word position-in-sentence, broken down by sentence length in the Dundee corpus. The panel headers show the sentence length. Dots represent mean values of single fixation duration. The gray area marks the 95% confidence interval. The solid lines are approximations of the data by the polynomial spline function with 3 knots.

Figure 3: Single Fixation Duration as a function of relative word position-in-line, broken down by line number on the screen in the Dundee corpus. The panel headers show the line numbers on the five-line screen. Dots represent mean values of single fixation duration. The gray area marks the 95% confidence interval. The solid lines are approximations of the data by the polynomial spline function with 3 knots.
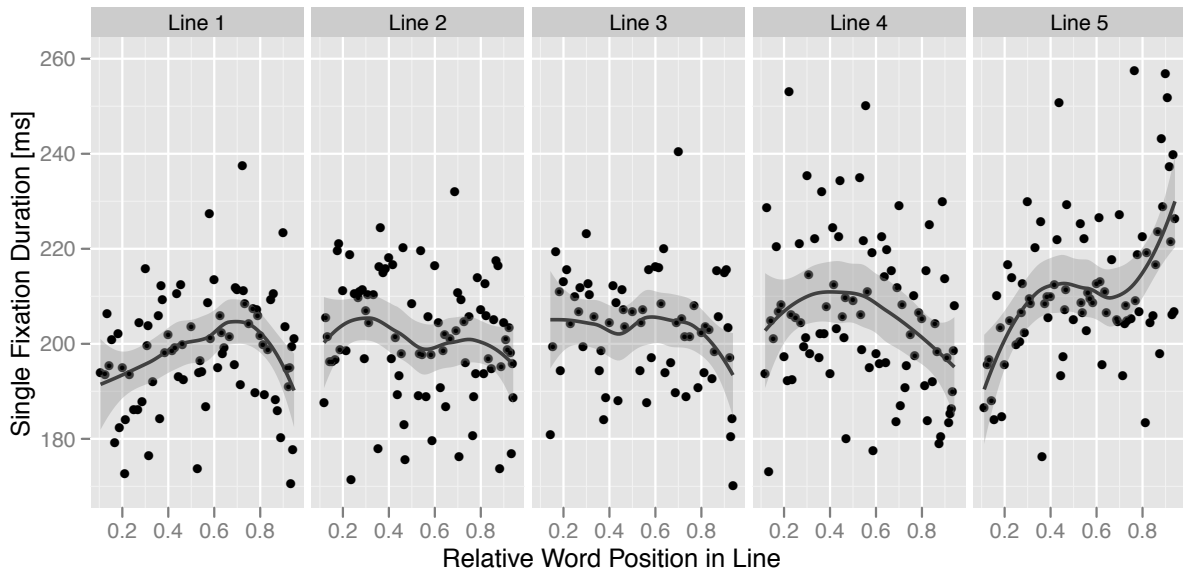
middle of those lines (by 5-15 ms), even though words in this region are not necessarily the ones that begin a sentence. Crucially, however, lines 1-4 demonstrate a *decrease* towards the end of the line. We note that the study of Pynte and Kennedy (2006) reported a sweeping non-linear increase in inspection times, yet their analyses collapsed word position-in-line across all lines and missed the finer structure of the effect. It is unclear what gives rise to the speed-up towards the end of the line. At present, we can only speculate that the speed-up is part of the typical processing of line breaks by the ocular system. Clearly, further research is necessary to examine this issue; for relevant evidence on how line breaks are treated by the oculomotor system in text reading, see Mitchell, Shen, Green, and Hodgson (2008).

Taken together, the patterns of results for word position-in-sentence and word position-in-line in the Dundee corpus provide a clear indication that the word position effect is bound to visual objects (e.g., line or screen), rather than syntactic or lexical objects (word, sentence or paragraph). Whether or not inspection times increase at extremes of the line is contingent on the line position on the screen, which supports our hypothesis that the two spatial regions are affected by different processing phenomena, the start-up and the wrap-up. The difference between lines also has important consequences for the notion of the wrap-up effect as we define it here. As lines on the screen are generally not aligned with sentence endings, it holds for all lines on the screen that the endings of those lines may lack typical indications of sentence ending: appropriate implicit prosody, punctuation marks, or the appropriate syntactic structure. And yet, the wrap-up effect is only evident in the screen's last line, just like it is evident at the end of a single-line sentence. What unites the screen's last line

and the single-line sentence? Possibly, it is the fact that they both precede a task which requires a mental representation of the available visual information: either an integration with the visual material on the next screen or a reply to a comprehension question. While translation of visual information into semantic codes has long been claimed as the underlying cause of the wrap-up effect for the clause-final word, we note that such translation needs not be confined to the end of a sentence or a clause. It may also occur at the end of a line (in single-line reading) or a screen (in paragraph reading). We now set out to establish what factors at what levels of processing give rise to the observed effects in sentence and in paragraph reading.

Causes of the POST effect: Lexical and contextual processing

In the Introduction we presented evidence for the correlations that word position-in-sentence shows with such characteristics of words as predictability and frequency. Since both predictability and frequency are major determinants of reading, it is crucial to investigate whether there is a unique influence of word position-in-text over and beyond the influences of these and other factors. We tackle this issue first for our data on sentence reading. For instance, the data patterns we attribute to the POST effect might emerge due to larger numbers of lower-frequency words or of less predictable words occurring towards the end of the sentence. These explanations are qualitatively unlikely, though, because of the *positive* correlations of word position-in-sentence and predictability (PSC: $r = 0.43, p < 0.01$, DEMONIC: $r = 0.01, p > 0.05$) and word position-in-sentence and frequency (PSC: $r = 0.05, p = 0.07$, DEMONIC: $r = 0.12, p < 0.01$, DMORPH: $r = 0.07, p = 0.05$). That is, sentence endings

are more likely to contain more predictable and higher-frequency words, which are generally processed faster (e.g., Boston, Hale, Kliegl *et al.*, 2008; Ehrlich & Rayner, 1981; Rayner, 1998) and hence would have led to a sentence-final speed-up, which is contrary to our findings. Similarly, words may tend to be longer the further into the sentence they are. Given that longer words are read slower, the word position effect might be a mere consequence of the word length distribution across the sentence. Yet, the correlations of word length with word position-in-sentence are negative across three corpora [PSC: $r = -0.21, p = 0.05$; DEMONIC: $r = -0.29, p < 0.01$; DMORPH: $r = -0.20, p = 0.01$], indicating that words at the end of the sentence tend to be shorter than those in the beginning, and thus they are expected to lead to a speed-up, and not the observed slow-down.

We further tested the independent contribution of the word position-in-sentence by including it in statistical models along with multiple other predictors. We tested whether there is variance uniquely explained by this factor when the influence of potentially confounding variables is regressed out from the data. We introduced relative word position and absolute word position, separately, into linear mixed models (implemented in the statistical software package R (2007)) with single fixation duration and gaze duration as dependent variables, as well as a number of fixed effects, and participant, word and sentence as random effects. These models were fitted to PSC, DEMONIC and DMORPH data sets. This yielded the total of 12 models (2 word position definitions x 2 measures of reading times x 3 data sets). For all models, the fixed effects included length, frequency and lexical status (function or content word) for words N-1, N and N+1, amplitudes of incoming and outgoing saccades, and – for PSC

and DEMONIC– logit estimates of predictability norms for word N, N-1 and N+1. Given the non-linear nature of the POST relation with inspection times, we modeled absolute and relative word positions as non-linear predictors with linear, quadratic and cubic components (approximated by cubic orthogonal polynomials using function `poly` implemented in the statistical software `R`, with degree of 3). Crucially, across all models, absolute and relative word positions showed significant effects ($ps < 0.05$) predicting an increase in inspection times, while multiple other predictors and between-items and between-participants variance were accounted for[2]. Due to space limitations, we only present the models for the three data sets with log single fixation duration as a dependent variable, and with relative word position-in-sentence as a non-linear predictor in Table 2; **the non-linear effects of the absolute and relative word position-in-sentence were replicated in the models fitted to log gaze duration as well**.

Some of predictors used in our models were correlated (e.g., length and frequency of words N-1, N and N+1): this collinearity may lead to inaccurate estimates of the model parameters associated with these predictors. To avoid collinearity, we orthogonalized pairs of predictors with correlation coefficients $r > 0.2$. Orthogonalization was achieved by using the residuals of the model in which word frequency is regressed against word length as estimates of frequency of words N-1, N and N+1. The residualized values were strongly correlated with the original values (all $rs > 0.7$) and had the advantage of being orthogonal to word length. Likewise, we orthogonalized relative word position from predictabilty of word N and, since the order of residualization is important, we also orthogonalized predictabilty

of word N from relative word position. We refitted the models reported in Table 2 with different sets of orthogonalized predictors substituting for original predictors. Crucially, the effects of word position-in-text that are in the focus of this study have not changed their direction, nor have they lost their statistical significance at the .05-level. We conclude that the word position in sentence explains variance over and beyond the currently known major lexical and contextual determinants of reading times, and hence its effect is unlikely to be an artefact of those confounds. For simplicity, Table 2 reports the outcomes of models with un-orthogonalized predictors.

INSERT TABLE 2 ABOUT HERE

It is also noteworthy that in many sentences of our corpora, word 2 stood for the copula verb (e.g., *is* as in *She is tall.*), that is, a very frequent and short lexeme, which might have been processed faster than subsequent longer and less frequent words, thus accounting for the sentence-initial increase in inspection times. **If present, the effects of the frequency and length of the verb would be particularly strong if we zoom in on the first few words of the sentence and they might override the effect of word position in sentence. Looking at the sentence as a whole may wash away these effects due to increased noise. To test this prediction,** we re-fitted our statistical models to the subsets of PSC, DEMONIC and DMORPH with only words 2-4. These models (not shown) replicated the positive correlation of word position with single fixation duration and gaze duration, while controlling for word frequency and length. Thus, the start-up effect is unlikely to be caused by the distribution of word lengths or frequencies over this critical

26

region.

On the level of sentence structure, processing costs have been argued to stem from (a) the amount of the intervening linguistic material between an incoming word and another term of the syntactic dependency to which that word attaches, and (b) the number of co-dependent terms in that intervening material (Gibson, 1998). Demberg and Keller (2008) found that in a limited number of cases these factors co-determined fixation durations for verbs and nouns in the Dundee Corpus. Since distances between words are confounded with word positions, the POST effect may be masking the effect of the cost of syntactic processing. We associated each word in our Dutch data sets DEMONIC and DMORPH with its distance from the head of the dependency to which that word attaches (the distance was defined as 0 for dependency heads). To this end, Dutch sentences were parsed using Alpino, the syntactic dependency parser for Dutch (van der Beek, Bouma, Malouf & Noord, 2002). In agreement with Demberg and Keller (2008), the distance between co-dependent terms was not a significant predictor of inspection times in our statistical models for DEMONIC and DMORPH (ps > 0.1), so we rule out this measure of syntactic complexity as a cause of the POST effect. To sum up, a broad spectrum of lexical and contextual predictors cannot account for either the start-up or the wrap-up effects on inspection times in sentence reading.

To tackle the paragraph reading data, we fitted linear mixed models to the Dundee data set with single fixation duration and gaze duration as dependent variables (models not shown). The two major changes that we applied to the structure of the fixed and random effects described above and in Table 2 were as follows: (a) Word position was defined as

27

(absolute or, separately, relative) word position in line, rather than in a sentence, and (b) line number on the screen was introduced into the model as a main effect and in interaction with word position. The interaction was significant ($p < 0.001$) and faithfully reproduced the pattern in Figure 3: the word position did not correlate significantly with inspection times in lines 1-4, and it correlated positively with word position in line 5. Again, statistical models indicated that the word position effect cannot be reduced to the influence of many, potentially confounding, predictors of eye-movement measures. We note, however, that our conclusions regarding word position effect on paragraph reading remain tentative since they are based on only one corpus, and need further replication across corpora and languages. We now proceed to considering the oculomotor level of processing implicated in reading.

Causes of the POST effect: oculomotor processing

In the present section, we extend our exploration of causes for the POST effect to low-level oculomotor factors: The position of a word on the screen might affect the conditions under which the ocular system of the reader operates. One possible oculomotor account of the POST effect plausibly relates to the evidence reported by Vitu et al. (2004) for isolated word reading that saccades tend to gravitate towards the center of the screen and eye-movements differ in amplitudes depending on whether their launching sites are at the edge or in the center of the screen. It might be that saccades are triggered faster as they tend towards the screen center, while their initiation is delayed when the saccades are heading away from the center toward the right edge of the monitor. Such a pattern would generate the effect of the word position-in-sentence that we observed.

We examined this and other oculomotor accounts using data from a z-string reading experiment. This experimental paradigm is conceptualized as an oculomotor control condition to normal reading as it approximates reading without lexical and post-lexical processing. It is based on the following logic: In both normal and z-string reading, eye-movements will be influenced by roughly the same visual and oculomotor factors. Therefore, if only low-level visuomotor variables and the properties of the oculomotor system mediate the POST effect, the z-string data should show a pattern qualitatively similar to the normal reading data. In contrast, if the POST effect was predominantly driven by cognitive processes triggered by lexical material, we should not observe the distinct pattern of start-up and wrap-up effects on fixation times in the z-string condition. To test these predictions, we reanalyzed the data from the z-string reading experiment based on the Potsdam Sentence Corpus (see PSC) (Nuthmann, Engbert, & Kliegl, 2007; Nuthmann & Engbert, 2009). The visual summary of results are provided in Figure 4.

INSERT FIGURE 4 APPROXIMATELY HERE

First of all, the data from 26 readers presented with the experimental stimuli of PSC (normal reading) replicate the pattern we observe for other corpora of sentence reading, that is, the POST effect emerges as a cubic parabola-shaped increase in fixation durations. Importantly, when readers scan meaningless z-strings, (absolute or relative) word position exerts no effect on single fixation duration. Neither the start-up effect nor the characteristic wrap-up effect was observed. Apparently, even if low-level visuo-oculomotor processes drive the POST effect, it is crucial that the lexical material be read for the effect to emerge:
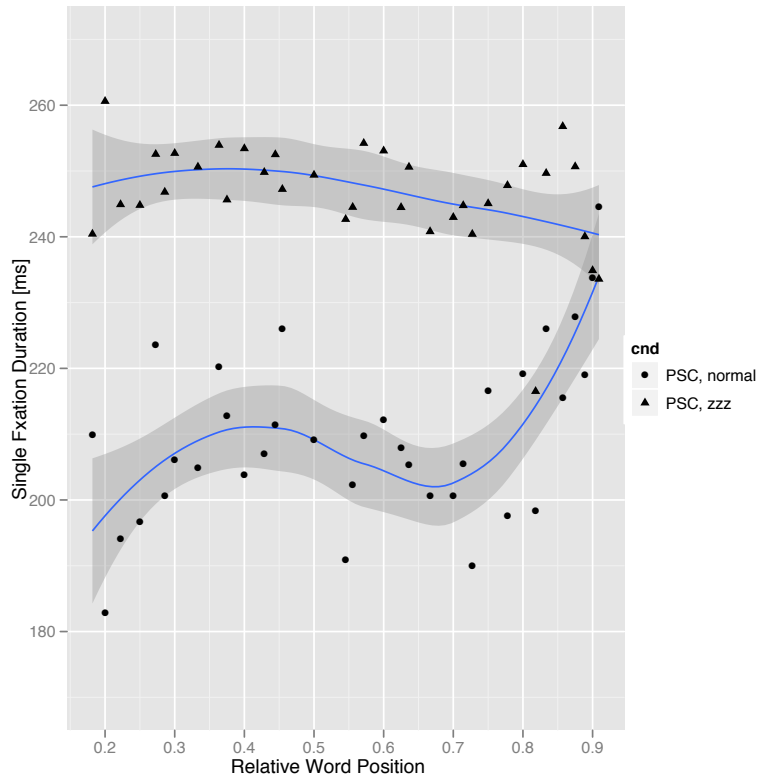
29

Figure 4: Single Fixation Duration as a function of relative word position in sentence, for readers of the Postdam Sentence Corpus (26 readers) and of matching z-strings (46 readers). The gray area marks the 95% confidence interval. The solid lines shows approximations of the data by the polynomial spline with 3 knots.

Scanning of strings of z-characters does not give rise to any variability in the processing cost as a function of string position in the line.

In line with the common routine of data preprocessing, sentence-initial and final fixations, as well as fixations on initial and final words were excluded from all data sets we considered so far. Yet fixation durations on word 1 in our sentences and lines are crucial, since they may co-determine the inspection times for subsequent words. If the inspection time for word 1 is long, it affords more parafoveal processing for word 2 and, possibly, word 3, thus leading to the reduced inspection times that we observe for those words. Our inspection of fixations on the first words of sentences in PSC, DEMONIC, DMORPH and z-string data sets ruled this explanation out. While fixations on the first words were on average substantially longer than those on the second words in PSC (about 40 ms) and the z-string data set (about 60 ms), they were shorter in DEMONIC (about 35 ms) and DMORPH (about 30 ms). The differences in data patterns most likely stem from the differences in experimental setups. As described above, the fixation point in DEMONIC and DMORPH was the location where the initial character of the first word of the sentence was presented, while first words of each sentence/string in PSC and the z-string reading experiment were displayed such that the eye was at the optimal viewing position. Hence, one possible explanation for the inflated durations of sentence-initial fixations in PSC and the z-string data is the robust Inverted Optimal Viewing Position effect (IOVP), such that the closer a fixation is to the optimal viewing position in the word, the longer it is in normal and mindless reading (cf. e.g., Nuth-mann & Engbert, 2009; Nuthmann, Engbert & Kliegl, 2007; Vitu, Lancelin & d'Unienville,

2007). This is unlikely, though, since there is no clear evidence for the IOVP effect for fixation durations at the initial word for any of our eye-movement data sets. Alternatively and more likely, it may take more time to detach a fixation when the eyes are in the center of a word (PSC and the z-string data) as compared to a setup, where stimuli appearing to the right of a fixation require an immediate saccade in the reading direction (DEMONIC and DMORPH).

A motor program of saccade planning in line reading

Yet another possibility for an underlying cause of the POST effect stems from the observation made by Pynte and Kennedy (2006). They point at the increase in the amplitude of incoming saccade over the entire line as a function of word position-in-line in English and French components of the Dundee corpus (Figure 2 in Pynte & Kennedy, 2006) and claim that "[t]he shape of the curves, with shorter saccades at the beginning and end of the line, points to the existence of a motor program modulating saccade planning over a whole line of text at a time." (Pynte & Kennedy, 2006, p. 3790).

It is logically possible that facing the task of reading a line of text, readers plan and launch an oculomotor program that aims at maintaining an optimal speed of moving through a line of text in reading for comprehension. Such a program would be sensitive to characteristics of line as a visual object (for instance, line length or spatial frequency information), and it would be superimposed on other oculomotor program(s) tied to the word-level processing. Perhaps the best way to envision the workings of the hypothesized line-level program is by way of metaphor. The eye-movement behaviour during reading of a line of text may be

akin to the behaviour of a sprint runner. The runner's speed is zero in the beginning of the track. The runner accelerates by making short frequent steps, which become longer and less frequent as the runner's speed increases and reaches its maximum (Mero, Komi, & Gregor, 1992). In this metaphor, the step length is analogous to saccade amplitude[3]. Thus, the global program of running the sprint at the optimal speed includes the acceleration and the constant speed phases, and it co-determines the measurable characteristics of the runner's footwork complementary to the local programs activated in planning each specific step.

If the runner's behavior metaphor is translatable into eye-movement measures, we expect readers to start off with short saccades and gradually increase their amplitude until - after a few words - they reach the optimal speed of reading. Shorter incoming saccades strongly predict shorter fixations in all our data sets (see regression coefficients in Table 2); also Kliegl et al. (2004) report the incoming saccade amplitude as the single strongest predictor of inspection times. Thus, the increase in saccade amplitude in the line-initial region would explain the start-up increase in reading times. The same might hold for the line-final region: if saccades increase in their amplitude at the end of the line, this would give an alternative, non-semantic, account for the observed wrap-up effect.

As our assumption about the line-level motor program hinges on the pattern of saccade amplitudes across lines of text, we considered this eye-movement measure as a function of the word's POST in our sentence-reading data. Crucially, in all three data sets there was a substantial increase (on the order of 1.5 - 2 chars) in the incoming saccade amplitude in the first 2-5 words, see Figure 5. The increase either carries with a much smaller magnitude
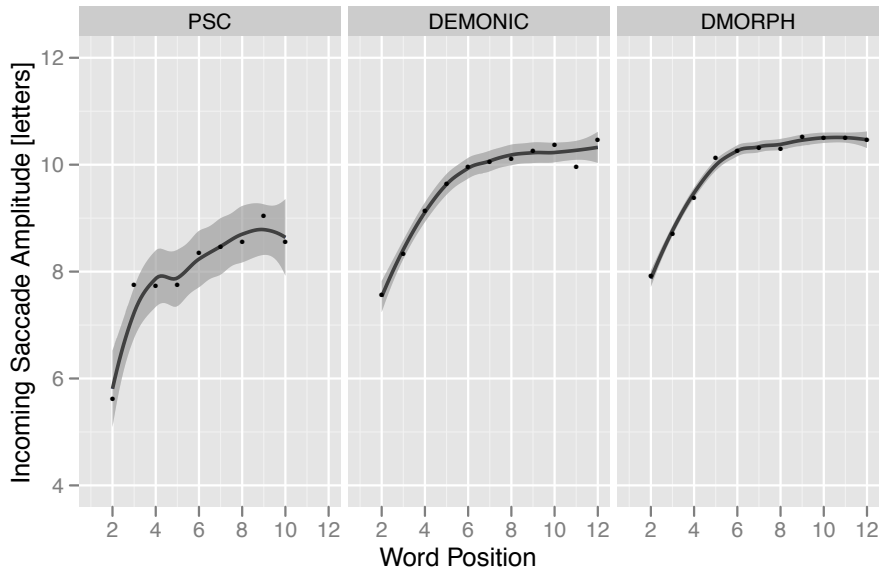
Figure 5: Incoming saccade amplitude (in characters) as a function of absolute word position-in-line, for PSC, DEMONIC and DMORPH.

over the rest of the sentence (PSC and DMORPH) or stays flat for the rest of the sentence (DEMONIC). The data on incoming saccade amplitudes are not available to us for the English or French components of the Dundee corpus, yet the patterns in Figure 6 in Pynte and Kennedy (2006) and the start-up increase in reading times in the English component of the corpus hint at a similar pattern in the paragraph reading as well. We speculate that the pattern of the line-initial increase in saccade amplitudes would emerge more clearly in English and French if it were broken down by the position of line on the screen, as done in Figure 3 for single fixation duration.

INSERT FIGURE 5 APPROXIMATELY HERE

The plots in Figure 5 provide an insight into the causes of components of the POST

effect. First, these behavioural patterns are compatible with the hypothesis of a line-level oculomotor program that allows readers to reach their optimal speed of moving through the line of text. We demonstrate that the start-up effect may arise due to the strategic saccade planning for the entire line: our examination of the start-up effect above also shows that it cannot be (fully) due to higher-level cognitive processes or to low-level oculomotor programs operating on the word-to-word basis. **It also appears that reading of the lexical material, rather than scanning of strings of symbols, is essential for whether the saccade planning program is set in action, as no start-up effect is observed in the scanning of z-strings.** Second, the patterns show no evidence of the qualitative change in saccade amplitudes in the line-final regions across our corpora (with the possible exception of French paragraph-reading, which we cannot presently examine in more detail). So, unlike the start-up effect, the oculomotor program that we advocate here cannot give rise to the wrap-up effect. This implies that the start-up and the wrap-up components of the observed POST effect appear to be of a different nature, with the former effect being a likely outcome of a saccade planning program that readers activate for the task of reading lines of text.

GENERAL DISCUSSION

We have shown strong and qualitatively similar effects of word position-in-text (POST) on reading times in three corpora of single-line sentences and a corpus of paragraph-based reading, in three languages. The effect has a functional form of a cubic parabola trend with a steeper inflation of inspection times in the beginning of single-line sentences and lines on the multi-line screen, as well as an increase in reading times at the end of single-line

35

sentences and the last line on the screen. The methodological consequence of our findings is that the POST effect needs to be accounted for in single-line and paragraph reading either by including it as a control in statistical models, or by avoiding the extremes of lines as regions for items that are factorially contrasted across conditions. The POST is independent of a number of factors which routinely emerge as co-determiners of reading behaviour, such as contextual predictability, word length, frequency, and lexical status, as well as syntactic complexity of a sentence. Moreover, the POST effect is unlikely to stem from the low-level word-to-word oculomotor processes implicated in reading, as suggested by the comparison of normal reading with the reading of z-strings: the latter revealed no change in inspection times across the string sequence. The POST effect is apparently a compound effect that reflects different processes (dubbed here as the *start-up* and the *wrap-up*) applying to the line-initial and line-final regions respectively. The different origins of the start-up and the wrap-up effects become evident in that the presence of the former, but not the latter, effect can be linked to the gradual increase in saccade amplitudes as the eyes proceed further into the line of text. We interpret the line-initial systematic patterns in saccade amplitudes as evidence for a strategic program of saccade planning which spans entire lines of text rather than enables transitions between specific pairs of words (cf. Pynte & Kennedy, 2006) **and is apparently specific to lexical reading rather than scanning of strings of symbols**. Such a program may be necessary to maintain the optimal (maximum) speed of reading over textual objects larger than isolated words. Refinement of parameters (other than the word position-in-text) that determine the operation of such a program is a topic for future

research; see below.

What explains the inflation of reading times at the end of a single-line sentence and of a screen-final line? Earlier work on the wrap-up effect, defined for the clause- or sentence-final word, advocates semantic integration of the sentence as its underlying cause, with the end of the clause being signaled by a variety of orthographic, visual, syntactic or prosodic cues (cf. Hirotani et al., 2006; Warren et al., 2009). The present data suggest that the increase in reading times is not confined to the last one or two words in a clause or a sentence, but rather increases incrementally over several words preceding the sentence/line boundary. A parsimonious hypothesis about the relation between earlier studies and our results would be that the inflation of reading times in the entire sentence-final region is caused by the same phenomena as the inflation on the last word in the sentence. Under this hypothesis, our finding of the wrap-up increase in inspection times in the **screen-final** lines of the Dundee corpus extends the range of contexts that are known to trigger the wrap-up effect. Notably, **screen-final lines** do not necessarily carry the marks of a sentence closure (punctuation marks, prosodic intonation or syntactic cues). Hence the effect observed at the end of single-line sentences and screen-final lines may be primarily triggered by the boundary of a visual unit, which signals the need for semantic integration of the present visual information. Interestingly, not every visual unit triggers a wrap-up effect. Lines that are followed by other lines on the screen do not display a wrap-up increase in inspection times towards the end of the line. The data suggest that the crucial factor in whether or not the wrap-up effect occurs at the end of a line of text is (a) whether this visual information will be available for further

perusal and/or (b) whether it needs to be **processed** – as a whole, or in an integrated form – to meet the comprehension test or to maintain the text processing even as it is interrupted by the change of the screen. Apparently, the completion of the visual uptake for a single-line or a screen can invoke the eye-movement behaviour similar to the one previously reported for the syntactic closure of the sentence or a clause.

The proposed account of the POST effect and its components generates a number of testable hypotheses for future research, outlined below:

1. Reading of sentences in which words are shuffled randomly will demonstrate the start-up effect, as readers will launch the oculomotor program to process the line of text at the optimal speed. Yet there will be no wrap-up effect, as such sentences would be semantically vacuous and impossible to integrate.

2. Scripts with the possibility of a vertical arrangement of characters (Chinese, Japanese) may show a start-up effect in the vertical dimension, with initial characters being processed faster than final ones. That would point at the independence of the effect from the physical (horizontal) position of a word/hieroglyph on the screen.

3. Reading of extended lines of text, exceeding the usual limit of about 80 characters imposed by the computer screen size and demands of visual acuity, may show a different pattern for the start-up effect. Just like the long-distance runners accelerate less than dash runners do in the very beginning of their run, we expect a weaker start-up effect in the reading of extended lines as compared to that effect in the beginning of shorter lines, like the ones explored in this study. Also, manipulation of spatial frequency of text lines is expected

to have an effect on the hypothesized line-level program of saccade planning and to correlate with the pattern of saccade amplitudes.

To conclude, the present investigation across corpora, languages and types of reading demonstrates for the first time that word position-in-text strongly influences word reading times in sentence and paragraph reading. The comparison of corpus data on sentence reading with the paragraph-based corpus further reveals that the POST effect emerges in the processing of visual objects (e.g. lines or paragraphs) rather than syntactic or linguistic objects (words or sentences). The POST effect appears to have two independent components: (a) the increase in inspection times over several sentence-initial words, arguably caused by the oculomotor program of saccade planning over the line of text (the start-up effect), and (b) the increase in inspection times over several sentence-final words, probably due to semantic integration and **comprehension** triggered, among other factors, by the boundary of a visual object like a line or a screen (the wrap-up effect). We suggest that word position-in-line become part of the standard battery of variables considered when analyzing the eye-movement record, along with word frequency, length or predictability. Also, we believe that current models of eye-movement control in reading will benefit from incorporating the line-level oculomotor program of saccade planning in their architectures.

ENDNOTES

[1]Here we expand the conventional use of the term *wrap-up* to apply to several sentence-final words, rather than just one clause- or sentence-final word (see our survey of the wrap-up literature above). For this reason, our empirical findings cannot be related directly to earlier

discussions of the phenomena occurring at the very last word of the clause.

[2]We also introduced absolute and relative word position-in-sentence, separately, in the model with over 80 predictors developed by Kliegl (2007) for PSC, with single fixation duration and gaze duration as dependent variables. Word position effect reached significance in those models as well. We also note that the three data sets differed somewhat in the structure of their effects, with the most prominent discrepancy being between the significant effects of predictability of words N-1, N and N+1 on single fixation duration for word N in the PSC dataset, and the weak, not significant effects of those predictors in DEMONIC. We attribute this discrepancy to the fact that 84% of words in DEMONIC were not guessed even once in the norming study, yielding the Cloze predictability score of 0; the analogous percentage in the PSC is a mere 33%. It is not surprising that a predictor that does not have much variability in our DEMONIC data fails to explain variance in the respective eye-movement record. At present, it is unclear whether the low predictability scores in DEMONIC were due to the web-based method of data collection, a lesser number of participants (50 complete protocols vs. 83 in the PSC norming study) or to the nature of the experimental sentences in this data set: we leave this question for further investigation.

[3] **To expand on the metaphor, we observe that the position of the shoe, and of the foot in the shoe, remains relatively stable during the contact with the ground, while the rest of the body is propelled forward from the position behind the foot. Similarly, the eyes stop completely at the word during fixation, modulo minor fluctuations, but the cognitive operations (including foveal and parafoveal visual**

uptake; spill-over processing of word N-1 and lexical access to word N; attention shift; or planning of the next saccade) are in progress. Notably, the step length and frequency are also dependent on the speed of the body, quality of the ground (lexical material under inspection), quality of the shoe (individual visual acuity), the physical ability of the athlete (reading proficiency or level of fatigue), the launching position of the preceding step (incoming saccade) and multiple other factors (Hunter, Marshall, & McNaire, 2004). Finally, the runner's speed is analogous to the distance (in characters) from the left screen edge divided by the total time – fixations plus saccades – that it took to run through this distance, which is not a measure commonly employed in eye-tracking studies, including the present one.

References

[1] Aaronson, D. and Ferres, S. (1983). Lexical categories and reading tasks. *Journal of Experimental Psychology. Human Perception and Performance*, 9:675–699.

[2] Aaronson, D. and Scarborough, H. (1976). Performance theories for sentence coding: Some quantitative evidence. *Journal of Experimental Psychology. Human Perception and Performance*, 2:56–70.

[3] Balogh, J., Zurif, E., Prather, P., Swinney, D., and Finkel, L. (1998). Gap-filling and end-

of-sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain and Language*, 61:169–182.

[4] Beek van der, L., Bouma, G., Malouf, R., and Van Noord, G. (2002). The Alpino Dependency Treebank. In Theune, M., Nijholt, A., and Hondorp, H., editors, *Computational Linguistics in the Netherlands CLIN 2001. Selected papers from the Twelfth CLIN Meeting.* Rodopi.

[5] Boston, M., Hale, J., Kliegl, R., Patil, U., and Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, 2:1–12.

[6] Chang, F. (1980). Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. *Memory and Cognition*, 8:58–64.

[7] Dambacher, M., Kliegl, R., Hofmann, M., and Jacobs, A. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research*, 1084(1):89–103.

[8] Demberg, V. and Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109(2):193–210.

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

[10] Ferreira, F. and Henderson, J. (1993). Reading processes during syntactic analysis and reanalysis. *Canadian Journal of Experimental Psychology*, 47:247–275.

[11] Gernsbacher, M. (1990). *Language comprehension as structure building.* Lawrence Erlbaum Associates, Hillside, NJ.

[12] Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1):1–76.

[13] Hill, R. and Murray, W. (2000). Commas and spaces: Effects of punctuation on eye movements in sentence processing. In Kennedy, A., Heller, D., and Pynte, J., editors, *Reading as a perceptual process*, pages 565–589. Elsevier, Amsterdam.

[14] Hirotani, M., Frazier, L., and Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *JML*, 54(3):425–443.

[15] Hunter, J., Marshall, R., and McNair, P. (2004). Interaction of step length and step rate during sprint running. *Medicine and science in sports and exercise*, 36(2):261–271.

[16] Just, M. and Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological review*, 87(4):329–54.

[17] Kennedy, A. and Pynte, J. (2005). Parafoveal-on-foveal effects in normal reading. *Vision Research*, 45:153–168.

[18] 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):8.

[19] Kliegl, R., Grabner, E., Rolfs, M., and Engbert, R. (2004). Length, frequency, and

predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, 16:262–284.

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

[21] Kuperman, V., Bertram, R., and Baayen, R. (in press). Processing trade-offs in the reading of Dutch derived words. *Journal of Memory and Language.*

[22] Mero, A., Komi, V., and Gregor, R. (1992). Biomechanics of sprint running. *Sports Medicine*, 13:376–392.

[23] Mitchell, D., Shen, X., Green, M., and Hodgson, T. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the Selective Reanalysis hypothesis. *Journal of Memory and Language*, 59(3):266–293.

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

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

[Patson and Warren] Patson, N. and Warren, T. Eye movements when reading implausible sentences: Investigating potential structural influences on semantic integration. *Quarterly Journal of Experimental Psychology*, in press.

[27] Pynte, J. and Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. *Vision Research*, 46(22):3786–3801.

[28] Pynte, J., New, B., and Kennedy, A. (2008a). A multiple regression analysis of syntactic and semantic influences in reading normal text. *Journal of Eye Movement Research*, 2:1–11.

[29] Pynte, J., New, B., and Kennedy, A. (2008b). On-line contextual influences during reading normal text: A multiple-regression analysis. *Vision Research*, 48(21):2172–2183.

[30] R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

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

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

[33] Rayner, K., Kambe, G., and Duffy, S. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology A*, 53(4):1061–1080.

[34] Rayner, K., Sereno, S., Morris, R., Schmauder, A., and Clifton, C. (1989). Eye move-

ments and on-line language comprehension processes. *Language and Cognitive Processes*, 4(3):21–49.

[35] Reichle, E., Warren, T., and 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.

[36] Van Petten, C. and Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition*, 18:380–393.

[37] Vitu, F., Kapoula, Z., Lancelin, D., and Lavigne, F. (2004). Eye movements in reading isolated words: Evidence for strong biases towards the center of the screen. *Vision Research*, 44(3):321–338.

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

[39] Vitu, F., O'Regan, J., Inhoff, A., and Topolski, R. (1995). Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception and Psychophysics*, 57(3):352–64.

[40] Warren, T., White, S., and Reichle, E. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and EZ Reader. *Cognition*, 111:132–137.

Table 2: Summary of regression models: random and fixed effects

		PSC			DEMONIC			DMORPH		
Random Effects			Variance	Std.Dev.		Variance	Std.Dev.		Variance	Std.Dev.
	Word ID	Intercept	0.007	0.083	Intercept	0.002	0.046	Intercept	0.003	0.053
	Subject ID	Intercept	0.018	0.132	Intercept	0.011	0.106	Intercept	0.000	0.019
	Sentence ID	Intercept	0.001	0.037	Intercept	0.001	0.025	Intercept	0.010	0.100
	Residual		0.060	0.245		0.097	0.312		0.089	0.298

		PSC			DEMONIC			DMORPH		
Fixed Effects		Estimate	Std. Error	t value	Estimate	Std. Error	t value	Estimate	Std. Error	t value
	Intercept	*5.311*	0.012	457.1	*5.512*	0.025	222.00	*5.440*	0.030	182.50
Word Position										
	Relative Position (linear)	-0.384	0.713	-0.5	*2.542*	0.546	4.65	*2.095*	0.498	4.20
	Relative Position (quadratic)	0.840	0.576	1.5	*2.105*	0.508	4.15	*1.212*	0.447	2.71
	Relative Position (cubic)	*2.860*	0.523	5.5	*2.386*	0.448	5.32	*1.546*	0.429	3.61
Present Word										
	Frequency n	*-0.040*	0.005	-8.3	*-0.012*	0.001	-11.24	*-0.010*	0.001	-8.13
	Length n	0.001	0.002	0.6	*0.002*	0.001	2.43	*0.003*	0.001	2.56
	Predictability n	*-0.017*	0.002	-7.6	-0.013	0.019	-0.66			
	Lexical Status	*0.100*	0.012	8.2	*0.018*	0.006	2.96	*-0.020*	0.007	-2.92
Past Word										
	Frequency n-1	*-0.016*	0.003	-5.8	-0.001	0.004	-0.31	*0.002*	0.001	2.10
	Length n-1	*-0.007*	0.001	-6.2	*0.003*	0.001	4.81	*0.003*	0.001	2.98
	Predictability n-1	*-0.008*	0.002	-3.4	-0.022	0.022	-1.02			
	Lexical Status n-1	*-0.026*	0.007	-3.8	-0.004	0.004	-1.00	-0.002	0.006	-0.38
Future Word										
	Frequency n+1	*-0.014*	0.003	-5.0	*-0.018*	0.004	-4.30	*-0.003*	0.001	-3.66
	Length n+1	*-0.011*	0.001	-11.1	-0.001	0.001	-1.81	-0.002	0.001	-1.22
	Predictability n+1	*0.004*	0.002	2.1	-0.001	0.020	-0.05			
	Lexical Status n+1	*-0.019*	0.007	-2.7	-0.007	0.004	-1.56	0.001	0.006	0.09
Viewing Position										
	Incoming Saccade Amplitude	*0.023*	0.000	51.2	*0.005*	0.000	22.75	*0.005*	0.000	16.77
	Outgoing Saccade Amplitude	*0.008*	0.001	14.4	*0.002*	0.000	13.25	*0.002*	0.000	6.25
	Viewing Pos in Word (linear)	*-0.110*	0.006	-17.2						
	Viewing Pos in Word (quadratic)	*-0.304*	0.018	-17.1						

We mark in italics the coefficient estimates that showed the p-values below 0.05. PSC stands for the Postdam Sentence corpus, DEMONIC for the Dutch Eye-Movements Online Internet Corpus and DMORPH for a large-scale experiment on reading of Dutch sentences.