A dynamic adjustment model of saccade lengths in reading for word-spaced orthographies: evidence from simulations and invisible boundary experiments

Jarkko Hautala, Stefan Hawelka, Otto Loberg & Paavo H.T. Leppänen

Published online: 08 Dec 2021.

Submit your article to this journal ⬈

View related articles ⬈

View Crossmark data ⬈

Routledge
Taylor & Francis Group

A dynamic adjustment model of saccade lengths in reading for word-spaced orthographies: evidence from simulations and invisible boundary experiments

Jarkko Hautala [a,e], Stefan Hawelka [b], Otto Loberg [c] and Paavo H.T. Leppänen [d,e]

aNiilo Mäki Institute, Jyväskylä, Finland; bCentre for Cognitive Neuroscience & Department of Psychology, University of Salzburg, Salzburg, Austria; cDepartment of Psychology, Bournemouth University, Bournemouth, UK; dDepartment of Psychology, University of Jyväskylä, Jyväskylä, Finland; eCentre for Interdisciplinary Brain Research, University of Jyväskylä, Jyväskylä, Finland

ABSTRACT

Contemporary models of eye movement control in reading assume a discrete target word selection process preceding saccade length computation, while the selection itself is assumed to be driven by word identification processes. However, a potentially more parsimonious, dynamic adjustment view allows both next word length and its content (e.g. orthographic) to modulate saccade length in a continuous manner. Based on a recently proposed center-based saccade length account a new regression model of forward saccade length is introduced and validated in a simulation study. Further, additional simulations and gaze-contingent invisible boundary experiments were used to study the cognitive mechanisms underlying skipping. Overall, the results support the plausibility of dynamic adjustment of saccade length in word-spaced orthographies. In the future, the present regression formula-based computational model will allow a straightforward implementation of influences of current and next word content (visual, orthographic, or contextual) on saccade length computation.

Research on eye movements during reading has provided a wealth of knowledge about the interplay among visual perception, cognition, and motor control, leading to highly complex and accurate computational models (Engbert et al., 2005; Reichle et al., 1999; Reichle & Sheridan, 2015; Reilly & Radach, 2006; Risse et al., 2014; Snell et al., 2018). The central phenomena that these models attempt to explain are "where" and "when" to move the eyes (e.g. Inhoff et al., 2003), the former being quantified as the direction and the length of a saccade, and the latter corresponding to fixation duration. In general, the prevalent models assume that forward saccade length is mostly programmed to enable and optimize word recognition within a single fixation (Reichle et al., 1999; see also Vitu, 2003). However, recent findings of a stronger influence of spatial width than the number of letters in saccade targeting during reading suggest

more involvement of visual constraints in saccade planning than has been generally assumed (Hautala et al., 2011; Hautala & Loberg, 2015; Hermena et al., 2017; Yao-N'Dré et al., 2014), providing a general motivation to revisit the "where" question. The prevalent models further assume that intended word skipping occurs when a word is parafoveally recognized to a sufficient degree, in which case saccade is targeted to a subsequent word. However, the tenability of this assumption is seldom formally tested (McConkie & Zola, 1984). We conducted such a formal assessment in the present study.

In the prevalent models, forward saccade length computation consists of the following main steps: First, activation dynamics of the current and the *parafoveal* words affect which word (n, n+1, or n+2) is selected as a saccade target. The activation processes themselves vary greatly among the models

and are out of the scope of the present paper, but generally do include both visual and lexical processes (an interested reader is referred to the above provided references). Second, saccade length is programmed to hit the center of the selected word, but then redirected—to some extent—towards the *preferred saccade length* of seven letter spaces (e.g. by a coefficient of 0.21 for a typical fixation duration of 215 ms in the E-Z Reader model; Reichle et al., 1999).

Word skipping results partially from the preferred saccade length pushing the eyes to "overshoot" the *n*+1 word at near launch distance (McConkie et al., 1988, 1989). In addition, skipping is assumed to occur when the next word *n*+1 is recognized rapidly enough from the parafovea, and consequently, the saccade is programmed to hit the center of word *n*+2. For example, according to the E-Z Reader model (e.g. Reichle et al., 2012), high-frequency spatial information from a parafoveal word is extracted rapidly and in parallel while lexically processing a foveal word. As soon as the foveal word has been recognized, attention shifts to the parafoveal word starting the first stage of lexical processing, the "*familiarity check*", which consists of orthographic processing leading to activation of a word form. Skipping occurs if this "familiarity check" is completed early enough to cancel the default saccade to the next word—thus, no full recognition of the next word is required. Previous simulations have shown that the model produces realistic, unique effects of word length, frequency, and predictability on skipping rates (Pollatsek et al., 2006), and that the speed of the "familiarity check" is the earliest word content-related process that substantially affects saccade lengths (Reichle et al., 2013).

However, some of the underlying assumptions of the above described *discrete control* view have also been criticized (e.g. Albrengues et al., 2019; Yang & Vitu, 2007; Yao-N'Dré et al., 2014). First, the idea of saccade correction toward the preferred saccade length (in other words, a *systematic error*) relies heavily on a single influential study (McConkie et al., 1988). This study, however, has a crucial shortcoming: The saccade length dynamics were studied only for interword saccades from word n to word *n* +1, therefore neglecting the influence of refixations and skipping saccades on the saccade length computation (see Yao-N'Dré et al., 2014). Reanalyzing McConkie et al.'s (1988) data, Reilly and O'Regan (1998) found a best fit for a model in which the

eyes were targeted to the longest word in a 20-letter perceptual window. However, such a mechanism is not psychologically plausible, as it would suggest that readers stereotypically skip even words of moderate length if the words are followed by longer words. Later, Vitu and colleagues thoroughly analyzed saccade control in reading, including refixation and skipping saccades, by predicting saccade length or landing position with launch distance to next word and its length (Albrengues et al., 2019; Yang & Vitu, 2007; Yao-N'Dré et al., 2014). The conclusion drawn from these studies is that readers tend to proceed with much more uniform saccade lengths than would be expected on the basis of interword saccades only.

Vitu argued that saccade lengths are *dynamically adjusted*, that is, in a continuous manner without the involvement of a target word selection process, even in word-spaced orthographies (Yang & Vitu, 2007). According to Vitu's *center of gravity account*, the eyes are targeted to the center of visual saliency when taking into account cortical receptive fields of human vision (Yao-N'Dré et al., 2014). However, no computational model has yet been provided to demonstrate the plausibility of this view. Dynamic adjustment models of saccade length have been developed for reading the unspaced Chinese orthography (Liu et al., 2018; see also Li & Pollatsek, 2020; Yu et al., 2020). According to these models, saccade lengths increase as a function of parafoveal preprocessing of the next word, which is facilitated by the predictability and frequency of the next word. However, because these models are limited to the reading of Chinese (which does not clearly demarcate word boundaries with blank spaces), they are not informative about the influence of word spaces on saccade control—the subject of the present investigation. However, there already exists direct empirical evidence that linguistic properties of upcoming words are also used at the very least to fine-tune interword saccade lengths in alphabetic orthographies (Albrengues et al., 2019; Bicknell et al., 2020; Hyönä et al., 2018; Radach et al., 2004).

Recently, a dynamic adjustment account of saccade lengths in word-spaced orthographies was suggested by Cutter et al. (2018), who showed that the mean saccade length is rapidly adapted to the current text and that over and above of a launch distance, the length of a currently fixated word also has a fundamental role in determining saccade lengths (see also Wei et al., 2013

for similar findings in Chinese). According to Cutter et al. (2018), the distance between the centers of adjacent words *n* and *n*+1 (i.e. *center distance,* CD) reflects an ideal saccade length, which *may* already be partly perceived when fixating a word *n*-1 and thus previewing the words *n* and *n*+1. This preliminary program may then be corrected to some extent by the actualized landing position (LD) to word *n*, that is, the actual launch distance to the beginning of word *n*+1. The authors labeled this view as a *center-based saccade length* (CBSL) account. The idea of the preparation of a saccade sequence is in contrast to the discrete control models, which assume that very early initiation of a new saccade plan cancels the current saccade plan. Saccade-targeting studies have provided support for both stances (see Becker & Jürgens, 1979; McSorley et al., 2019). To the best of our knowledge, direct evidence for the preparation of saccade sequences in reading has not yet been provided, so we remain agnostic on this issue.

Overall, the fundamental issue of whether a reader's eye movements in word-spaced orthographies are dynamically or discretely controlled requires a revisit. Therefore, we conducted a simulation study (Study I) that introduced and validated a novel computational implementation of the CBSL against observed data and the E-Z Reader model (e.g. Reichle & Sheridan, 2015), which assumed the target word selection process. Then, in an experimental study (Study II), we focused on word skipping, which is the decisive phenomena for the discrete vs. dynamic control issue. First, crucial predictions from the E-Z Reader and CBSL models were derived from simulations, which were then empirically tested with gaze-contingent invisible boundary experiments.

Study I

In this study, we investigated whether a novel computational model of the CBSL could produce the benchmark word length effects of a shift of landing position toward the word beginning, increase refixation probability, and decrease skipping probability. As an extension to Cutter et al. (2018), we hypothesized that *minimum saccade length* would ensure that at least some new visual information would be sampled during each fixation, whereas the *maximum saccade length* would prevent exceeding human visuo-attentional capabilities (Hautala et al., 2011; Hautala & Loberg,

2015; Hermena et al., 2017; Yao-N'Dré et al., 2014). In effect, we assumed that the ability of the CBSL to target the next word center would produce realistic landing positions and that the minimum saccade length would push saccades beyond word *n*+1 from near launch distance and produce skipping. We further reasoned that influence of maximum saccade would be diverse: It could cause the landing position to shift towards word beginning for long words, and cause refixating of long words. The refixations, in turn, would bring the eyes close to the next word, and thus cause skipping.

The CBSL model was constructed by first estimating a regression formula from a dataset with the standard linear mixed model (LMM) analysis, which also provided optimized parameter values for the model. Non-linear terms in regression analysis allowed us to estimate the minimum and maximum saccade lengths; that is, saccade length does not increase even if CD or LD continues to increase. The model was then supplemented with a standard *random oculomotor error*.

In external validation (see Faber & Rajko, 2007), the model was used to simulate saccade lengths on a sentence corpus for which the model was not initially trained, and for which there was also observed eye movement data available (Schilling et al., 1998). Thus, a statistically equally good or better fit to observed data (see Ludden et al., 1994) for the CBSL model relative to the benchmark of the E-Z Reader model would provide a "proof-of-concept" for the dynamic adjustment mechanism of saccade targeting in word-based orthographies (and not just, e.g. the reading of Chinese; Liu et al., 2018; Yu et al., 2020).

Methods

Empirical datasets

Three datasets from different languages were used in the present study, with the underlying assumption that saccade targeting in reading of word-spaced orthographies follows universal principles (Liversedge et al., 2016). For example, in a cross-linguistic study, a similar effect of average word length in a sentence on the number of fixations was observed for English and Finnish reading (Liversedge et al., 2016). In line with the universality assumption, Figure 1 shows highly similar relative landing position (fixation position in a word divided by its length in character spaces), skipping,

and refixation probabilities as a function of word length among Finnish (Hautala & Loberg, 2015), English (Schilling et al., 1998), and Russian datasets (Laurinavichyute et al., 2019). Concerning the missing relative landing positions in English, a previous study by Plummer and Rayner (2012) reported even lower (by <0.1) relative landing positions for English (0.465 for 4.7 letter words, and 0.397 for 8.3 letter words) than observed here in Finnish and Russian datasets.

The simulated skipping probability and refixation probabilities were compared with the observed means derived from Schilling et al. (1998), kindly provided by Erik Reichle. In this study, fluent English readers read 48 sentences. This sentence corpus and observed data were the basis for developing the E-Z Reader model and optimizing its parameters, and thus, they provided a maximally strict reference for evaluating the CBSL model predictions. Because no word length specific landing position data for the Schilling dataset were available, the observed relative landing position values were taken from the published Russian corpus of a single sentence reading experiment of fluent readers (Laurinavichyute et al., 2019).

Center-Based saccade length model and simulations

The regression formula of the CBSL model was derived by fitting a linear mixed model to an eye movement dataset of a single sentence reading experiment in Finnish (Hautala & Loberg, 2015). In this experiment, participants read 160 sentences including 1260 words with a mean length of 7.8 letters ($SD = 2.9$). The subset of data from 23 fluent readers was the subject of the present analysis. The first and last words in a sentence were excluded from the current analyses. The experiment included

a manipulation in which half of the sentences were presented with doubled word spaces in a counterbalanced manner. Because these extra word spaces directly influence saccade lengths, these trials were excluded from the present analysis. Further, Cutter et al. (2017, 2018) showed that center distance between adjacent words has an immediate effect on saccade length computation, so carryover effects of the extra spaces on normal sentences are of no concern.

As the stimulus sentences were presented in the proportional Calibri font (i.e. variable letter widths), the pixel saccade length values were divided by the mean letter width of 10 pixels to derive saccade length estimates in letters. Saccade lengths were calculated as the difference between adjacent fixation positions, and only forward shifts of fewer than three words were subject to linear mixed modeling. Extreme saccade lengths of 300 pixels were excluded (31), as well as launch distances of over 150 pixels (14). These criteria led to 12,181 observations with the following descriptive information expressed in letters: $M = 6.95$, $SD = 2.69$, $Min = 0.09$, and $Max = 28.3$. The overall first-pass measures of landing position, refixation, and skipping probability means reported in Figure 1 were exported from BeGaze software (SensoMotoric Instruments, Teltow, Germany).

To estimate the regression model for predicting saccade lengths, the data were subjected to LMM analysis with lmer-package in R (Bates et al., 2019). Here, the saccade length was predicted by intercept, CD, LD, and their non-linear, squared terms. In contrast to convention, non-centered values of independent variables were used because these estimates could then be used directly in the CBSL simulation model. No interaction terms among the factors were included due to them causing substantial collinearity problems, as diagnosed with
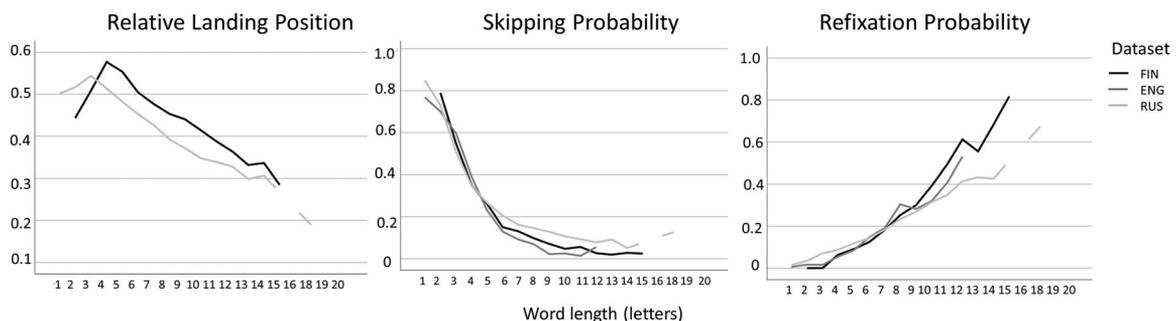

Figure 1. Empirical means of dependent eye movement measures in different datasets.
Note: FIN = Finnish, ENG = English, RUS = Russian.

variance inflation indexes derived with the vif-function of the car-package in R (Fox et al., 2007). The fixed factors were defined with poly-function (R Core Team, 2002), which was set to return the estimates of orthogonal polynomials in the original scale. Random intercepts of participants and items were included. Predicted means were plotted with the *emmeans* package (Lenth, 2018).

All of the fixed effects were significant, therefore confirming the importance of also including the non-linear terms in the CBSL model: Intercept ($\beta = 3.85$, $SE = 0.37$, $t = 10.4$, $p < .001$), CD ($\beta = 0.54$, $SE = 0.10$, $t = 5.7$, $p < .001$), CD^2 ($\beta = -0.019$, $SE = 0.007$, $t = -2.67$, $p = .008$), LD ($\beta = 0.32$, $SE = 0.03$, $t = 11.8$, $p < .001$), and LD^2 ($\beta = -0.032$, $SE = 0.003$, $t = -12.3$, $p < .001$). These beta estimates led to a regression formula (Eq. 1). Random effect variances were 0.14 ($SD = 0.37$) for items and 0.83 ($SD = 0.91$) for participants, and residual variance was 5.96 ($SD = 2.44$).

$$\text{Saccade length} = 3.85$$
$$+ 0.54*CD - 0.019*CD^2$$
$$+ 0.32*LD - 0.032*LD^2 \quad (1)$$

With the help of the estimated marginal means presented in Figure 2, the following interpretations were suggested for each term in the model. The minimum planned saccade length is mainly determined by the intercept as well as the influence of minimum realistic values of CD (2) and LD (0). The asymptote of planned saccade length (a maximum) is mainly determined by the non-linear effect of CD: an increase from 11 to 15 letter has almost no effect on saccade length. The non-linear effect of LD induces a clear shortening of a saccade at a very near launch distance, presumably reflecting an attempt to secure sufficient visual perception of word n+1. Saccade shortening is even stronger at a far launch distance (–0.43 per letter), which is possible only for long words. This effect presumably reflects the need to secure visual perception for an end of the currently fixated long word. Within these boundaries, saccade length increases by ~0.5 letter spaces for every letter increase in CD.

The simulation for the CBSL model was conducted by using Eq. (1), supplemented with the random oculomotor error component identical to the one used in the E-Z Reader. To provide a strict test for the CBSL model, the simulations were run on the Schilling corpus only. A standard Monte Carlo simulation scheme with 1,000 runs was adopted. The first and last words of each sentence were excluded when calculating means.

Simulations with the E-Z Reader model

The simulation with the E-Z Reader-java applet was run for the Schilling corpus with default parameter values without simulation of interword regressions (http://www.erikdreichle.com/downloads.html). To confirm the importance of lexical processing on eye movement control in E-Z Reader, we also ran simulations with the model's saccade targeting formulas only (E-Z saccade). The reader is referred to, for example, McGowan and Reichle (2018) for a detailed account of the E-Z Reader model and Supplementary file for method details.

Analysis of model fits

First, descriptive results of the CBSL were provided. Then, to compare the simulation results among the models, the difference between observed and predicted values, that is, errors, were calculated for each model, with the Russian dataset being the observed reference for landing position analysis and the English dataset for the refixation and skipping probability analyses. The item-specific means of absolute error values for the three models (CBSL, E-Z Reader, and E-Z saccade) were subject to one-way analysis of variance (ANOVA) and multiple comparison with Bonferroni adjustment. The null hypothesis was that the prediction error is equal across the models.

Results

Simulation results

The diagnostic plots in Figure 3 show that the CBSL model reproduced the classical landing position dependencies on word length, launch distance, and normal distribution of saccade lengths, indicating that the model is based on a realistic mechanism of eye movement control.

In Figure 4, the upper panel shows observed and simulated means for relative landing position, skipping, and refixation probabilities as a function of word length, while the lower panels show more closely the prediction errors with inter-item variability. Univariate ANOVA indicated highly significant differences in the goodness-of-fit among the models for landing position (Welch's $F(2, 804) =$
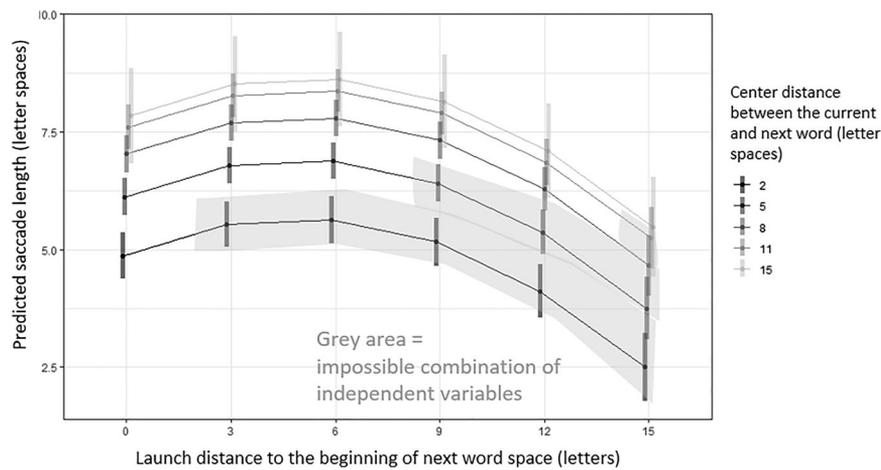
Figure 2. Estimated marginal means of the linear mixed model across varying center and launch distances.

Note: Grey area approximates the impossible predicted values. For example, for a minimum center distance of two, both the previous and next word can cover only two letter spaces, whereas for a center distance of 15, the previous word can cover up to 28 letter spaces (14 + 1).
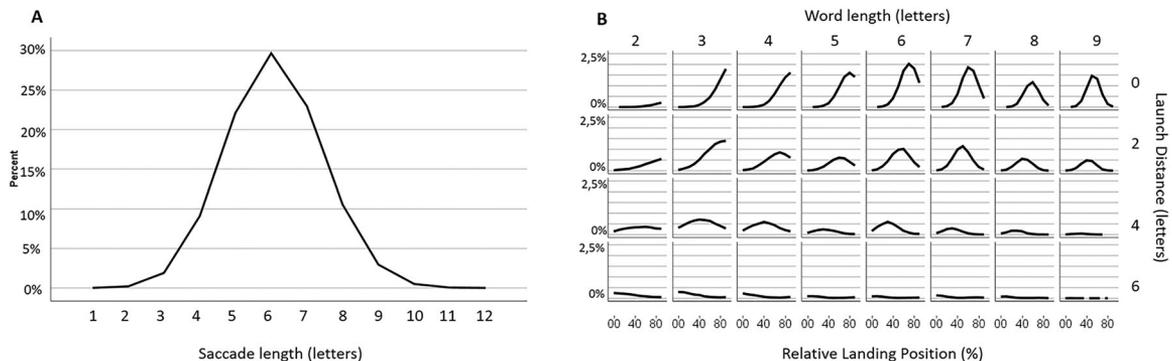

Figure 3. Diagnostic information of CBSL simulation on ENG corpus.

Note: Panel A shows a histogram of saccade lengths. Panel B shows landing position distribution as a function of word length and launch distance to the beginning of the next word.

169.9, $p < .001$), skipping (Welch's $F(2, 860) = 26.1$, $p < .001$), and refixation probability (Welch's $F(2, 870) = 144.2$, $p < .001$). For landing position, multiple comparison with Bonferroni correction specified that the CBSL ($M = 0.074$, $SD = 0.08$) produced better fit than the E-Z saccade ($M = 0.099$, $SD = 0.04$) and E-Z Reader models ($M = 0.16$, $SD = 0.07$), ps $< .001$. The qualitative inspection of the model fits in Figure 4 indicates that all models systematically overshot the observed landing position. This is especially true given the even more leftward relative landing position values observed in a previous study in English (Plummer & Rayner, 2012). Note that this result is not surprising given that E-Z Reader's landing position predictions have been previously validated only qualitatively. What is more important is that CBSL produces realistic

landing position predictions for the entire range of word lengths.

For skipping, multiple comparisons indicated equal fit for the CBSL ($M = .175$, $SD = 0.15$) and E-Z Reader ($M = .191$, $SD = 0.17$) and clearly inferior fit for the E-Z saccade model ($M = .266$, $SD = 0.22$). Although the curves in Figure 4 suggest that the E-Z Reader excels in predicting skipping probabilities, Figure 5 shows how the small average error is actually due to the substantial underestimation of skipping in a subset of sentences. The large underestimation of skipping by the E-Z saccade model confirms the expected importance of lexical processes in the generation of skipping within the E-Z Reader framework. In a similar vein, CBSL steadily predicts less skipping for all word lengths, suggesting the need to
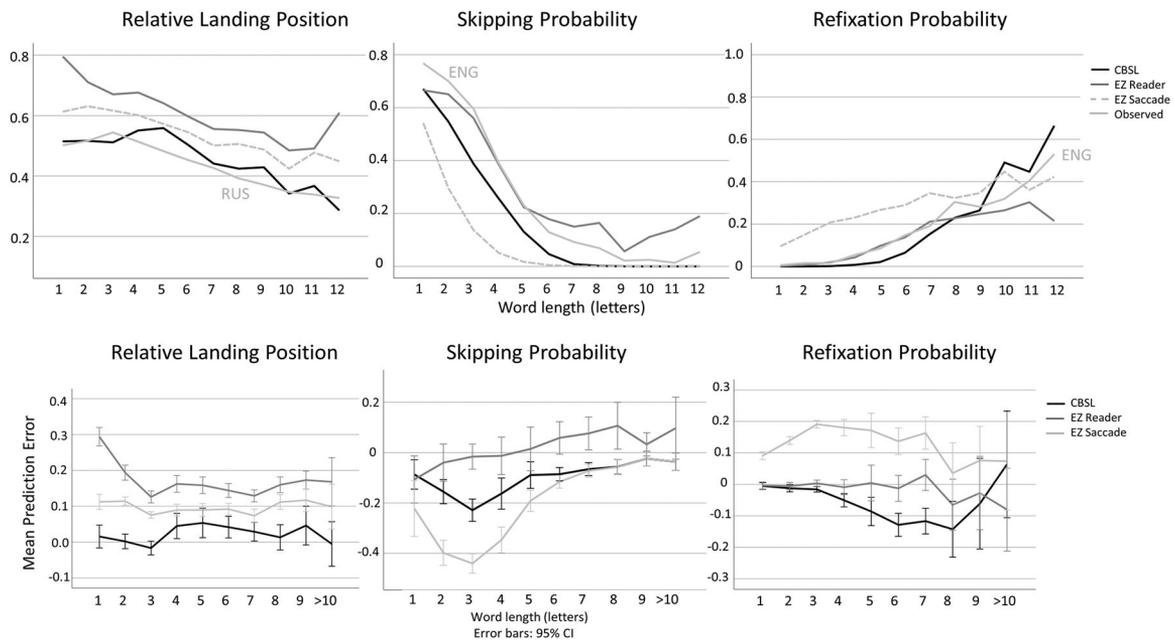
Figure 4. Simulation results.

Note: Upper panels show the observed reference values and simulation results presented as means for relative landing position, skipping probability, and refixation probability for varying word lengths. Lower panels show the prediction errors with item variability shown with 95% confidence intervals (CIs). Negative values indicate more rightward landing position values or more skipping or refixation in observed data than predicted by the models (underestimation), and vice versa. Statistically, the prediction error is null when the CIs cross zero.

implement lexical processes to modulate saccade lengths to some extent.

For refixation probability, multiple comparisons indicated equal fit for CBSL ($M = .081$, $SD = 0.12$) and E-Z Reader ($M = .079$, $SD = 0.10$) and clearly inferior fit for the E-Z saccade model ($M = .182$, $SD = 0.10$). The CBSL underestimated refixation probability in short words and overestimated it for the longest words. The former suggests the need to implement a corrective refixation mechanism, whereas the latter may stem from a lexical process to "skip" endings of long words whose endings are predictable on the basis of the already perceived/processed beginning. The large overestimation of refixations by the E-Z saccade model alone confirms that the lexical processes of the E-Z Reader model strongly inhibit refixations for already recognized words, leading to highly accurate predictions.

Summary of study I findings

Analysis of saccade length provided support for the concepts of minimum (4.9 letters) and maximum (8.5 letters) saccade lengths in reading. Up to CD
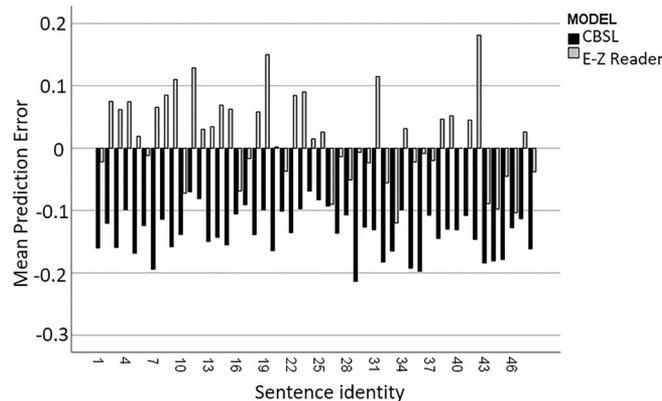

Figure 5. Between-sentence variability of prediction error for skipping probability in the E-Z Reader and CBSL simulations.

values of 11 letters, the saccade length increased by 0.5 letters for every letter increase. Saccades were longest at an LD of six letters. At shorter LDs, targeting of upcoming word reduced saccade length. Instead, the reduction of saccade length at LDs farther than six letters presumably reflects the need to refixate the end of long words. The statistical comparisons indicated a good fit for the CBSL in predicting observed landing positions and equally accurate predictions for skipping and refixation probability when compared with E-Z Reader's predictions. Yet, the current CBSL model predicted too high of a rate of refixations on long words relative to observed data, which may indicate that lexical processes are required to inhibit refixations on a well-recognized foveal word.

Study II

To study whether word skipping is based on the discrete or dynamic control mechanisms, we will first present simulated landing positions distributions across word boundaries for the discrete (E-Z Reader) and dynamic control (CBSL) models. The derived predictions are then empirically tested with eye-movement experiments designed to maximize skipping of target words, while ensuring that landing position distribution over successive words can be separated. The invisible boundary technique is utilized to study to which extent skipping is guided by fine-grained visual and orthographic information extracted from parafoveal perception of upcoming word.

Due to its ability to target the center of either word n+1 or n+2, the discrete control model (E-Z Reader) predicted a bimodal landing position distribution, whose peaks corresponded to the centers of these successive words (Figure 6(a)). However, in the dynamic model (CBSL), the saccade length is always computed according to word n and n+1 lengths only, resulting in a uniform landing position distribution over these words (Figure 6(b)). We were able to find only one previous study testing the bimodality -hypothesis, where McConkie and Zola (1984) reported a bimodal saccade length distribution, but only at very near launch distance to the next word.

In addition, E-Z Reader predicted that launch distance (Figure 6(c)) and frequency (Figure 6(d)) of word n+1 modulates the probability of targeting the center of n+1 or n+2 words. Although the lexical processes have not yet been implemented into the CBSL model, in principle, the content of the upcoming word would only be used to fine-tune saccade lengths.

Furthermore, both the CBSL and E-Z Reader models assume that many word skips are "overshoots". In the present study, the skipping rate of invalid previews may be taken as an approximation up to which extent skips of unpredictable short words are based on low-level visual information such as word spaces. However, the strong effect of a word's spatial width on skipping (Hautala et al., 2011; Hautala & Loberg, 2015; Hermena et al., 2017; Yao-N'Dré et al., 2014) raises the possibility that skipping occurs when a visual perception of sufficient quality has been sampled from the upcoming word. Accordingly, because vision gradient degrades toward larger eccentricities, odd visual information located at n+1 word beginning vs. end should be more readily detected and should affect skipping more. Previous invisible boundary studies with exterior letter replacement (Briihl & Inhoff, 1995) or partial visual degradation manipulations (Gagl et al., 2014) have not supported this possibility; however, the manipulations used in these studies also led to inflated foveal fixation durations, suggesting the manipulations interfered with higher-level orthographic and lexical processing (see also Choi & Gordon, 2014; Drieghe, 2008; Gordon et al., 2013; Plummer & Rayner, 2012; Reingold & Rayner, 2006). However, visually manipulated previews, such as faint fonts, can be restricted in the early stages of word recognition (see Drieghe, 2008; Reingold & Rayner, 2006), having no consequences on later foveal processing (Warrington et al., 2018).

In the present study, we developed a novel visual dot preview manipulation, in which the target words contained a subtle visual extra dot located either at the beginning or end of the item. If the dots inhibit skipping to a similar degree, that would suggest that the parafoveal perception has been subject to attentional enhancement early enough to affect saccade programming (for a review, see Li et al., 2016). Finally, if the dots affect skipping to a similar degree as nonword previews, then that would mean that preview-content guided word skipping depends on a process that is equally interfered by both visual dots and letter replacements, with a likely candidate being the activation of orthographic word representations. Such a result would also suggest a rather limited role for higher level lexical processing in skipping (Choi & Gordon, 2014; Gordon et al., 2013).
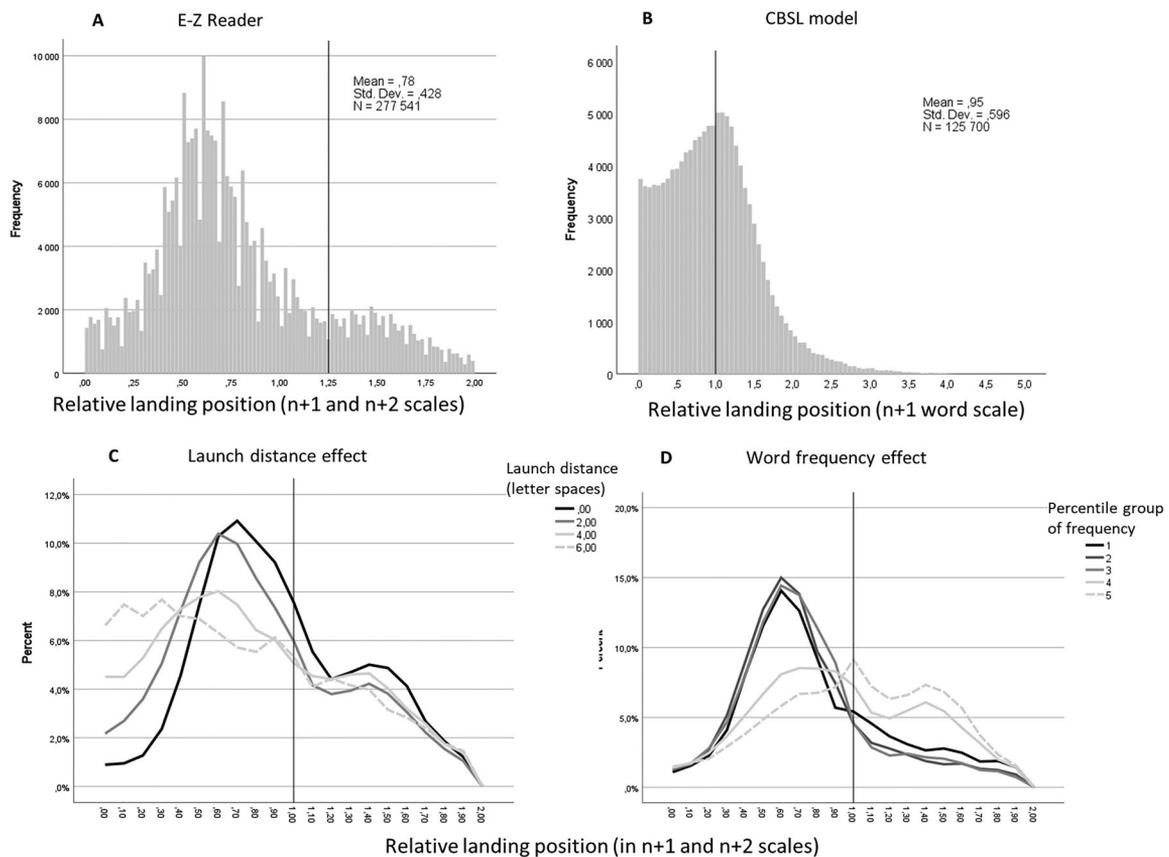
Figure 6. Simulated landing position distributions.

Note: Panel A: Histogram of simulated first-pass relative landing position over words n+1 and n+2 of any length for the E-Z Reader model. According to Hartigan's (1985) dip test, the overall distribution is not unimodal, D = .0107, p < .001. Panel B: Histogram of simulated landing position over words n+1 and n+2 for the CBSL model, when the target words (≤4 letters) were highly prone to skipping. The distribution is unimodal, D = .0005, p = 1. Panel (c–d): Simulated relative landing position distributions for E-Z Reader model at high skipping probability condition, when the length of word n+1 has five or fewer letters. To better tease out the bimodal peaks, the word n+2 was set to contain at least three letters. An increase in launch distance shifted the distribution peaks toward the left (Panel C), while a high word frequency of word n+1 increases the probability of targeting the center of word n+1 or n+2 (Panel D).

Methods

This study was conducted in three successive experiments (Exp 1, 2, and 3) sharing the majority of methodological details with each other (see Supplementary file).

Participants

There were 20 native Finnish-speaking participants in each of the three experiments (total N = 60), recruited via student organization mailing lists. The invitation criteria were uncorrected near vision and no reading disability. Table 1 provides descriptive information about the participants for each experiment. The work was conducted in accordance with the APA's and World Medical Association Declaration of Helsinki ethical standards in the treatment of sample, human or animal, or to describe the details of treatment. The research was approved by the Ethical Committee of the [Anonymized for Review]. The participants signed an informed consent form prior to the study and received a movie ticket as a reward for their participation.

Stimuli

There were 96 spatially narrow nominative target words (e.g. *ilo* [joy] and *tiili* [brick]) of 3–5 letters, and varying word frequency (M = 43, SDI = 110, min–max 0.2–713 occurrences in a million words). For reference, the mean word frequency in a Finnish psycholinguistic corpus is 44, calculated among words of higher frequency than one in a million (Huovilainen, 2018). A near-zero cloze probability (0.012 within a sample of six people) prevented guessing of target words.

The target words were previewed either with intact or invalid nonword or visual dot masks. Nonword previews were constructed by replacing a single letter of a target word with a visually resembling letter of equal width (e.g. *k-h, i-l, n-u, s-o, a-ä*) in letter positions 1–3, with the mean position being 1.7. The visual dot manipulations were drawn manually with image processing software. In the context of a narrow exterior letter, the dot was embedded into a letter space, and in the context of a wide exterior letter, such as *a*, *o*, *u*, or *n*, the dot was embedded inside the letter.

For Exps 2 and 3, the size of the dots was increased by adding gray pixels around the black four-pixel dot (Figure 7). To balance the probability of intact and invalid previews, the nonword preview condition was dropped for Experiment 3.

Apparatus

The EyeLink 1000 with tabletop mount (SR Research Ltd., Kanata, ON, Canada) was used to measure participants' eye movements. Eye dominance was determined by the distance hole-in-the-paper test (see Rice et al., 2008), and only the dominant eye was tracked. The stimuli were shown on an LCD screen (Asus VG-236, 1920 × 1080, 120 Hz, 52 × 29 cm; ASUSTek Computer Inc., Taipei, Taiwan) with a 61 cm viewing distance (one pixel corresponding to 0.2646 mm and 0.02485 visual degrees horizontally). Horizontal eye movements were recorded with a sampling rate of 2000 Hz,

Table 1. Descriptive information of participants in each experiment.

Measure		Experiment 1	Experiment 2	Experiment 3
Gender	(male/female)	3/17	6/14	4/16
Age in years	mean (SD)	25.5 (5.1)	25.6 (4.7)	24.2 (4.4)
Eye dominance	left/right	7/13	8/12	8/12
Comprehension accuracy	mean % (SD)	79 (11)	86 (9)	81 (15)
	min %	58*	71	50*
	max %	92	100	100
Sentence reading time (s)	mean (SD)	6.5 (1.4)	7.0 (1.7)	7.5 (3.1)

Note: [a]The six low-comprehending participants (<60% accuracy) nonetheless read the sentences properly, with their mean sentence reading times ranging from 4.9–8.5 s. One participant in each experiment became aware of the preview manipulations, but their data was kept in the analyses due to the Latin square rotation of the preview conditions across target words. Exclusion of their data did not change the overall pattern of the results, with the exception that the effect of the preview condition on first-pass skipping was no longer significant in Exp 1 ($p = 0.11$).

and participants' heads were stabilized with chin and forehead rests. Default three-point horizontal calibration covering the entire screen width was conducted prior to each experiment, with a mean calibration error of 0.14 degrees. The calibration points were presented on the same single line as the experimental sentences.

Procedure

Text was shown centered on the screen with proportional Arial 16-point font. The horizontal width of one letter was by average 3.1 mm, which corresponds to 0.291 visual degrees and 11.7 pixels at the middle of the screen. The width of a word space was 3 mm. Although some sentences covered almost the entire screen width, the target words were always located within the central 30 degrees of the screen.

To prevent sentence and word properties from producing effects among the four experimental conditions, each target word received every type of parafoveal preview according to a Latin square rotation of the preview conditions, with the restriction that the same type of preview was not presented for both the first and the second target word in a sentence. Thus, each participant in an experiment received a unique assignment of preview conditions for target words.

To maximize the skipping of target word n, it was embedded into a sentence between longer words n-1 and n+1. Due to refixations, the long n-1 word provided a near launch distance to word n, whereas a long word n+1 helped in determining whether a skipping saccade was targeted to it. To prevent fatigue, the number of sentences was kept low (48), and every stimulus sentence contained two target words. To engage readers in thorough reading, comprehension statements focusing on target word meaning were presented for a random 25% of the sentences. Participants responded (with mouse buttons) to these statements with an 81.4% accuracy rate (Table 1).

Two practice trials were presented at the beginning of the experiment, with one of them followed by a comprehension statement. To see a stimulus sentence, participants fixated on a circle appearing on the left side of the screen. The fixation was validated by the experimenter using a key press. If the validation was not successful, a recalibration was conducted.[*] After reading the sentence, participants pressed the left mouse
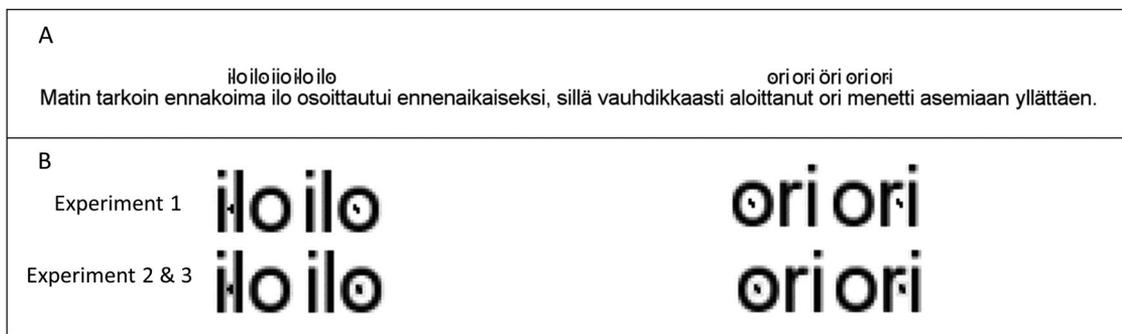
Figure 7. Example of a stimulus sentence.

Note: Panel A: The invalid previews for target words ilo (joy) and ori (stallion) are shown in their original size above the sentence. The sentence translates as follows: "Matti's carefully anticipated joy turned out to be premature, as the fast-started stallion suddenly lost its position". Panel B: Visual dot preview masks of target words magnified in four.

button to continue on to the next sentence or to receive a comprehension statement. The participant and the experimenter were situated in separate rooms.

Eye movement data processing

Fixation and saccade data were exported by SR Data Viewer software with a default saccade detection threshold of 30 deg/s. The functioning of the boundary triggers was inspected using the fixation and saccade data. The technical performance of the triggers was acceptable, with an average synchronization delay for screen changes of 6.08 ms, with a 2.3 ms *SD*. There were only four occurrences with >10 ms delays in a sample of 10 datasets. Data for target words with premature or delayed trigger firings were excluded from the analyses (see Supplementary file).

Analyses

First, experiment-specific LMM analyses were run. Among these results only the ones relevant to answering the research questions are provided in the Results -section (see Supplementary file for the full analyses). Then, Hartigan's (1985) dip test for unimodality was run for landing position distributions, using the dip test R package (Mächler, 2013). Because bimodality was not observed, there was no support for the discrete control view. Therefore, the dynamic adjustment mechanism was

studied further in a post-hoc analysis. Here the effective invalid preview conditions (nonword in Exp 1, large dots in Exp 3) were contrasted with the intact previews (intact and small dots in Exp 1, intact in Exp 3) to predict relative landing position in target word scale. The LMM also included factors of launch distance, and previous fixation duration. The random structure included intercepts and slopes for preview type for both the participants and items. Estimation of correlations between random effects was omitted by using the Afex package (Singmann et al., 2018). Landing position, previous fixation duration, and launch distance values greater than five *SD* from the mean were excluded.

Results

In Exp 1, there was a main effect of the preview condition, $\chi^2(3) = 10.8$, $p = .013$, for the probability of additional fixation measure,[†] resulting from the significant contrasts between the intact ($M = 6\%$) and the nonword ($M = 12\%$) conditions ($OR = 0.48$, $SE = 0.13$, $z = -2.72$, $p = .033$) and the dot_begin ($M = 6\%$) and the nonword conditions ($OR = 0.48$, $SE = 0.13$, $z = -2.69$, $p = .036$). In Exp 3 with large dot and intact trials, there were no differences between the conditions for the next fixation duration, $\chi^2(3) = 2.17$, $p = .34$, or probability of additional fixation, $\chi^2(3) = 1.27$, $p = .53$. Thus, the visual dot previews had no effect on later foveal processing measures, indicating that the visual dot

---

[*]The high calibration accuracy may have been compromised to some extent by the usage of lengthy sentences and not employing a 3-point validation before each trial.

[†]The foveal processing effect did not manifest immediately on a fixation duration on a target word, but in a delayed manner as an additional fixation made to a target word. Presumably, this was due to using visually similar replacement letters (Marcet & Perea, 2018). Also, the present study was not optimized for studying foveal effects due to high skipping rate, which reduced the number of first-pass fixations.
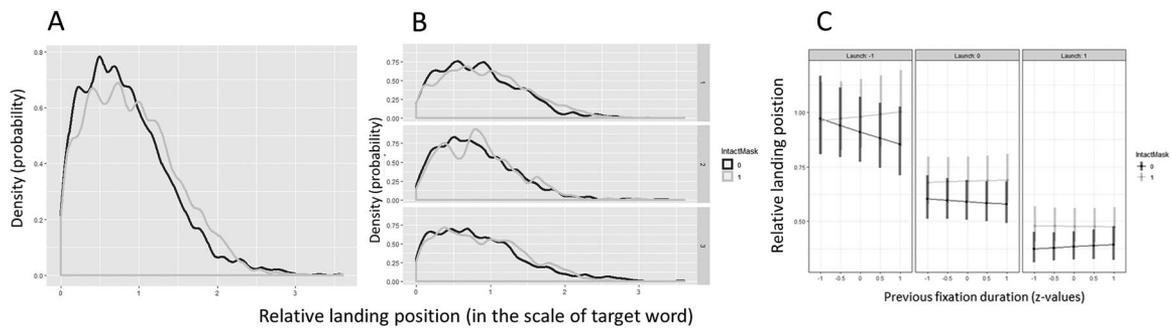
Figure 8. Results of the landing position analyses.

Note: Panel A: Density distributions of relative landing position for full dataset. Values 0–1 on the x-axis belong to the target word and higher values to its next word, including word space. Panel B: Density distributions broken down into different experiments and conditions. Panel C: Estimated marginal means with 95% CIs for a non-significant interaction of launch distance (standardized z-value), previous fixation duration (z-value), and preview type (0 = intact preview, 1 = invalid preview). The z-values −1, 0, 1 correspond to launch distances of 0.71, 1.41, and 2.13 visual degrees and previous fixation durations of 158, 208, and 276 ms, respectively.

manipulation exclusively affected early visual processing while preserving lexical processing.

In relation to intact previews, both nonword (Exp 1, $\chi^2(3) = 8.15$, $p = .043$) and large dot previews (Exp 3, $\chi^2(2) = 8.81$, $p = 0.01$) reduced first-pass skipping from the 0.4 level to the 0.3 level. Thus, up to 75% of skipping seems to be produced by the oculomotor control mechanism. Notably, the location of the visual dot had no effect on first-pass skipping. In Exp 3, the main effect of condition $\chi^2(2) = 8.81$, $p = 0.01$ resulted from a significant contrast between the intact ($M = 41\%$) and dot_begin ($M = 31\%$) conditions ($OR = 1.50$, $SE = 0.24$, $z = 2.61$, $p = .025$) and from the nearly significant intact and dot_end ($M = 33\%$) conditions ($OR = 1.41$, $SE = 0.22$, $z = 2.19$, $p = 0.07$). These results support rejecting the view that saccade length is adjusted in an online manner according to the quality of visual perception of an upcoming word, and they further provide support for the view that parafoveal vision is attentionally enhanced early enough to affect saccade planning. Moreover, the visual dot previews (irrespective of their location) in Exp 3 and nonword previews in Exp 1 reduced skipping equally ($z = -0.19$, $p = 0.98$) potentially indicating that both manipulations interfered with the activation of orthographic word representations. The result also speaks for a rather limited role of higher level lexical processing in skipping.

Hartigan's dip tests of landing position distributions indicated that the distribution was strongly unimodal (Figure 8(a)). For the relative landing position measured in the target word scale, the test results for whole data were $D = 0.007$, $p = 0.97$, for experiment-specific data, ps > 0.93, and for intact

and invalid conditions within experiments, ps > 0.47. In a scale of pixel distance to end of target word, the test results for whole data were $D = 0.004$, $p = 0.97$, for experiment-specific data, ps > 0.55, and for intact and invalid conditions within experiments, ps > 0.36. On a scale of relative landing position of word n+1 (values 0–1) and n+2 (values 1–2), the test results for whole data were $D = 0.007$, $p = 0.18$, for experiment-specific data, ps > 0.11), and for intact and invalid conditions within experiments, ps > 0.28). By visual inspection, there was some sign of bimodality for intact trials in Exp 2 (Figure 8(b)), but this pattern was not significant.

According to the full factorial LMM results, landing position shifts left as a function of increasing launch distance ($\beta = -0.43$, $SE = 0.03$, $t = -13.4$, $p < .001$), as well as after the presentation of an invalid preview ($\beta = 0.15$, $SE = 0.04$, $t = 3.34$, $p = .002$). The interaction between launch distance and preview ($\beta = 0.07$, $SE = 0.04$, $t = -1.82$, $p = .070$) approached significance, yet the effect was too weak to be interpreted with confidence. The stable influence of preview type, independent of launch distances or previous fixation durations (Figure 8(c)), suggests that preview perception-based saccade modulation is not directly dependent on vision gradient or previous fixation duration.

General discussion

The present results provide clear support for a dynamic oculomotor control mechanism being in play in the reading of word-spaced orthographies. Study 1 showed that the dynamic control model

(CBSL; Cutter et al., 2017, 2018) accurately predicted empirically observed landing positions and skipping and refixation probabilities across word lengths when compared with the benchmark discrete control model of E-Z Reader. Study II then added how a dynamic view predicted a unimodal landing position distribution over word n+1 and n+2 in high skipping probability conditions, while a discrete control model predicted a bimodal distribution. Further, simulations confirmed E-Z Reader's ability to predict that parafoveal information modulates the probability of targeting either of these words, while previous studies have shown how parafoveal word content is also used to fine-tune saccade lengths (Albrengues et al., 2019; Hyönä et al., 2018; Radach et al., 2004; White & Liversedge, 2006). Experimental data provided support for the dynamic adjustment (Li & Pollatsek, 2020; Liu, Huang, Gao, et al., 2017; Liu, Huang, Li, et al., 2017; Liu et al., 2018; Yu et al., 2020) view by revealing a unimodal landing position distribution with the content of the preview word only fine-tuning the landing position. However, before coming to a definite conclusion, more studies attempting to induce the bimodal landing positions need to be conducted (McConkie & Zola, 1984).

The results of Study 1 provided support for the concepts of minimum (~4.9 letters) and maximum saccade length (~8.5 letters) as a feasible alternative to the contested (Cutter et al., 2017, 2018) preferred saccade length of seven letter spaces. However, more research is needed to resolve to what extent the limits in planned saccade length result from visual and attentional constraints. There are several recent findings suggesting that the visual extent of words instead of the number of letters actually drives saccade length computation in reading (Hautala et al., 2011; Hautala & Loberg, 2015; Hermena et al., 2017; Yao-N'Dré et al., 2014). Results from Yao-N'Dré et al. (2014) provided evidence for visual constraints as increasing font size led to shorter saccade length in terms of number of letters. On the other hand, there is a vast amount of evidence that forward saccade length and perceptual span increase through reading development (e.g. Häikiö et al., 2009; Rayner et al., 2010). It may be that in very fluent reading, the attentional capacity approaches visual acuity constraints for letter encoding.

We found that saccade lengths were longest at a launch distance of six letters and then shorter at nearer and farther launch distances. At this distance, the saccade length was not strongly reduced by the targeting of the upcoming word, nor did the current word need to be refixated. The reduction of saccade length at far launch distances provides support for the view that refixations are partly produced by a devoted cognitive mechanism (McDonald, 2006). However, it is unlikely that the gradual mechanism provided by CBSL to explain refixations is entirely a realistic one, as it predicts most refixation to land on the very end of the word, which is not optimal for word recognition. The current CBSL model predicted too high a rate of refixations on long words relative to observed data, suggesting that refixation on long words may be omitted when the word has been already recognized on the basis of the first fixation. This leads to the conclusion that the cognitive mechanism underlying the determination of refixation saccade lengths requires further research.

Within the aforementioned constraints, the saccade length adjusts dynamically up to ~75% according to word center distance and launch distance. This order of magnitude suggests that readers of word-spaced orthographies heavily target word centers within the limits discussed above. The present results seem therefore to be more in line with the prevalent word-based models (Engbert et al., 2005; Reichle et al., 1999; Reichle & Sheridan, 2015; Reilly & Radach, 2006; Risse et al., 2014; Snell et al., 2018) than with the center of gravity view, suggesting that readers prefer rather constant saccade length in reading (Albrengues et al., 2019; Yang & Vitu, 2007; Yao-N'Dré et al., 2014). Cutter et al. (2017, 2018) suggested that readers will even preplan a saccade based on parafoveal preview of next and subsequent word, which would explain why current word length also directly affects saccade length independent from launch distance. Presumably, the reason for this preplanning is the significant time constraints affecting saccade planning, as it is estimated that accurate planning of saccades takes about 150 ms to complete (Jacobs, 1987). Although such a mechanism is certainly a realistic option, the exact time course of saccade planning in reading is not yet known and requires more research.

It was also found that the parafoveal percept-guided saccade length modulation is not dependent on launch distance or previous fixation duration (Eskenazi & Folk, 2015; Rayner et al., 2011;

White, 2007). This finding is in contrast with some of the discrete control model predictions and empirical findings that long fixation durations facilitate direct cognitive control of saccades (Khan et al., 2017; McConkie & Yang, 2003). The E-Z Reader model (Reichle et al., 1999) predicts inflated fixation durations prior to skipping due to saccadic preplanning, whereas parallel attention models predict that processing of the next word may affect fixation duration prior to skipping (Engbert et al., 2005; Reilly & Radach, 2006; Snell et al., 2018). Thus, the present results favor the view that parafoveal perception induces a time-independent influence on saccade lengths (Brothers et al., 2017; Schotter et al., 2012). One intervening factor in this respect may be the shorter fixation durations in transparent than opaque orthographies, such as in the Finnish language studied here (Liversedge et al., 2016), which may allow less time for parafoveal preview processes to exert cognitive control on foveal processing and saccade planning (see McConkie & Yang, 2003).

A substantial portion of skips (25%) was estimated to result from modulation of saccade length according to the content of the previewed word. This modulation was not contingent on the quality of visual perception of the previewed word, in which case an extra visual dot at the word beginning vs. the end would have affected skipping more. Therefore, it can be concluded that attentional processes enhance the visual perception of the upcoming short word early enough to affect saccade programming (Inhoff et al., 2005; Inhoff & Radach, 2014). Further, because nonword previews had an effect similar to that of visual dot previews, it can be concluded that preview content-guided skipping does not *necessarily* require lexical activation (see Veldre et al., 2020), but possibly depends on the activation level of orthographic word representation, with which the extra visual dots may have interfered (Choi & Gordon, 2014; Drieghe, 2008; Gordon et al., 2013; Plummer & Rayner, 2012; Warrington et al., 2018). However, the majority of the skips (up to 75%) were estimated to occur independently of preview content, and thus resulted from the oculomotor control mechanism (as formalized in the CBSL model) utilizing only word spaces. It can therefore be concluded that saccade lengths are predominantly computed according to *a priori* constraints of the reader's general visuo-attentional capabilities and coarse visual information of word spaces.

Invalid previews and word skipping typically have consequences on later foveal processing. After fixing the saccadic program, the processing of a parafoveal word continues, and its visual perception is further enhanced by the presaccadic attention mechanism (Li et al., 2016; Rolfs & Carrasco, 2012). In line with previous studies utilizing faint font manipulation (Glaholt et al., 2014; Reingold & Rayner, 2006; Warrington et al., 2018; White & Staub, 2012), but in contrast to partial visual degradation manipulations (Gagl et al., 2014), our subtle visual dot preview was not associated with foveal processing difficulties in the form of refixations, which was observed for our nonword previews (Choi & Gordon, 2014; Gordon et al., 2013; Hohenstein & Kliegl, 2014; Plummer & Rayner, 2012; Reingold & Rayner, 2006; Schotter et al., 2018, 2019). Thus, the task-irrelevant visual dots affected only the early level of lexical processing of words, which was overcome rapidly during later foveal processing (Warrington et al., 2018). Visual dot manipulation seems to thus provide a purer method to study early preview processes than visual degradation manipulation, which also has consequences on foveal processing (Gagl et al., 2014).

Study II also has several notable limitations. A possible caveat with visual manipulations is that in some circumstances, they may become saccade targets (or non-targets; see Reingold et al., 2016) themselves. We did not find such an effect; yet skipping was overall reduced in Exp 2, in which 75% of target words contained detectable anomalies. It is possible that when almost every sentence contains something weird, readers start to invest their attentional resources in detecting these anomalies by trying to fixate every word (for the strategic control of skipping, see Wotschack & Kliegl, 2013). Thus, it is recommended that the number of intact and visually invalid trials should be balanced in this type of research.

Importantly, the present findings may be rather specific to short, unpredictable, and not highly frequent words and therefore likely to represent skipping driven by the visuo-oculomotor mechanism. However, the skipping of highly predictable, frequent, and/or long words is likely to be driven more by higher-level cognitive processes, perhaps by a genuine discrete control mechanism (McConkie & Zola, 1984). In fact, there is evidence that word predictability has an equal influence on word skipping in short and long words (Rayner

et al., 2011), meaning that the effect is relatively greater in long words than in short words. Moreover, readers can also strategically adjust their reading, for example, by skimming parts of a text or skipping redundant words or entire parts of sentences.

The current work is a first step towards developing a formal computational model based on the CBSL view of saccade length computation. In continuing to explore the question of where to program saccades, a straightforward next step would be to model the influence of the current and next word's main lexical properties, such as frequency and predictability, on forward saccade length. Another limitation of this study concerns corrective refixations, particularly those after landing to the very beginning of relatively short words, which may be within the scope of the linear regression technique adopted here. However, regressive saccades may require the appliance of some rule-based algorithm as implemented in the E-Z Reader, in which refixation probability increases linearly as a distance of landing position from the word center. However, empirical data suggest that the refixation probability after late landing positions remained quite low in the majority of studies (Gagl et al., 2014; Hyönä & Bertram, 2011; McDonald & Shillcock, 2005; Nuthmann et al., 2005; Rayner et al., 1996), suggesting a quite limited role of these corrective refixations in determining saccade dynamics in reading.

Modeling the temporal dynamics of saccade programming would require more extensive reworking of the model. As already discussed, first, it needs to be resolved whether the preparation of a saccade from word n to n+1 already begins when fixating a word n-1. If this is the case, then omitting a word space between words n+2 and n+3 should have some detrimental effects on eye movement parameters, which corresponds with the estimates of fluent adults' perceptual span of 14–15 letter spaces to the right of the fixation (Rayner, 1986). It is also known that an increase in the character size of a moving window from 11 to 15 letters to the right of a fixation point is associated with an increase in forward saccade length, without any effect on average fixation duration (Veldre & Andrews, 2014). Thus, it seems that readers compensate for the lack of n+2 word length information by reducing saccade length instead of spending longer time programming it. From a modeling perspective, this preplanning may be implemented as a simple saccade plan accumulator, so that the preview of word n+2 length reduces the time needed to complete the saccade program when fixating word n+1. Such a mechanism would then predict the cost for programming the next saccade after skipping, because word n+2 length would not parafoveally coded. In line with this prediction, empirical data suggest inflated fixation durations by 10–20 ms after skipping (Kliegl & Engbert, 2005), which may be due to saccadic reprogramming.

Taken together, the present simulation results of the center-based saccade length account (Cutter et al., 2018) show that forward saccade lengths in the reading of word-spaced orthographies can also be explained to a high degree with a dynamic adjustment account, as previously shown in the non-spaced Chinese orthography (e.g. Liu et al., 2018). This adjustment process seems to consist of at least minimum and maximum saccade lengths, an "ideal" saccade length corresponding to the distance between the center of word n+1 and n+2, and online correction based on landing position on word n+1. These findings suggest a need to revisit the basic oculomotor control mechanisms in prevalent models of eye movement control in reading. The most realistic mechanism would probably allow both dynamic and discrete control (Liu et al., 2019), the latter coming more into play if the crucial parafoveal information is coarse and thus extracted very early (Reingold et al., 2016). Finally, the current CBSL model has a very limited scope in relation to comprehensive models of eye movement control. It might also be a worthwhile endeavor to study whether the saccade targeting mechanisms of CBSL could be integrated with the prevalent models (Engbert et al., 2005; Reichle et al., 1999; Reichle & Sheridan, 2015; Reilly & Radach, 2006; Risse et al., 2014; Snell et al., 2018).

Acknowledgements

Disclosure statement

Funding

Data availability statement

Simulated data that support the findings of this study are available from the corresponding author, JH, upon reasonable request. Simulation model is available at OSF repository: https://osf.io/f6u5p/?view_only=c67ad30e36df4548bf38b80ebcf3de7a Due to lack of written permissions from participants, empirical data is not shared publicly.

ORCID

Jarkko Hautala 🔟 http://orcid.org/0000-0002-7402-6364
Stefan Hawelka 🔟 http://orcid.org/0000-0002-5304-8521
Otto Loberg 🔟 http://orcid.org/0000-0001-5012-0818
Paavo H.T. Leppänen 🔟 http://orcid.org/0000-0002-8941-2225

References

Albrengues, C., Lavigne, F., Aguilar, C., Castet, E., & Vitu, F. (2019). Linguistic processes do not beat visuo-motor constraints, but they modulate where the eyes move regardless of word boundaries: Evidence against word-based eye-movement control during reading. *PLOS ONE*, *14*(7), 1–47. . https://doi.org/10.1371/journal.pone.0219666

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2019). *Lme4: Linear mixed effects models using Eigen and S4. R package v. 1.1-21. R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research*, *19*(9), 967–983. https://doi.org/10.1016/0042-6989(79)90222-0

Bicknell, K., Levy, R., & Rayner, K. (2020). Ongoing cognitive processing influences precise eye-movement targets in reading. *Psychological Science*, *31*(4), 351–362. https://doi.org/10.1177/0956797620901766

Briihl, D., & Inhoff, A. W. (1995). Integrating information across fixations during reading: The use of orthographic bodies and of exterior letters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 55–67. https://doi.org/10.1037/0278-7393.21.1.55

Brothers, T., Hoversten, L. J., & Traxler, M. J. (2017). Looking back on reading ahead: No evidence for lexical parafoveal-on-foveal effects. *Journal of Memory and Language*, *96*, 9–22. https://doi.org/10.1016/j.jml.2017.04.001

Choi, W., & Gordon, P. C. (2014). Word skipping during sentence reading: Effects of lexicality on parafoveal processing. *Attention, Perception, & Psychophysics*, *76*(1), 201–213. https://doi.org/10.3758/s13414-013-0494-1

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2017). Reading sentences of uniform word length: Evidence for the adaptation of the preferred saccade length during reading. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(11), 1895–1911. https://doi.org/10.1037/xhp0000416

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2018). Reading sentences of uniform word length–II: Very rapid adaptation of the preferred saccade length. *Psychonomic Bulletin & Review*, *25*(4), 1435–1440. https://doi.org/10.3758/s13423-018-1473-2

Drieghe, D. (2008). Foveal processing and word skipping during reading. *Psychonomic Bulletin & Review*, *15*(4), 856–860. https://doi.org/10.3758/PBR.15.4.856

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777–813. https://doi.org/10.1037/0033-295X.112.4.777

Eskenazi, M. A., & Folk, J. R. (2015). Skipped words and fixated words are processed differently during reading. *Psychonomic Bulletin & Review*, *22*(2), 537–542. https://doi.org/10.3758/s13423-014-0682-6

Faber, N. M., & Rajko, R. (2007). How to avoid over-fitting in multivariate calibration—The conventional validation approach and an alternative. *Analytica Chimica Acta*, *595*(1–2), 98–106. https://doi.org/10.1016/j.aca.2007.05.030

Fox, J., Friendly, G. G., Graves, S., Heiberger, R., Monette, G., Nilsson, H., … Suggests, M. A. S. S. (2007). *The car package. R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Gagl, B., Hawelka, S., Richlan, F., Schuster, S., & Hutzler, F. (2014). Parafoveal preprocessing in reading revisited: Evidence from a novel preview manipulation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(2), 588–595. https://doi.org/10.1037/a0034408

Glaholt, M. G., Rayner, K., & Reingold, E. M. (2014). A rapid effect of stimulus quality on the durations of individual fixations during reading. *Visual Cognition*, *22*(3–4), 377–389. https://doi.org/10.1080/13506285.2014.891542

Gordon, P. C., Plummer, P., & Choi, W. (2013). See before you jump: Full recognition of parafoveal words precedes skips during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 633–641. https://doi.org/10.1037/a0028881

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, *102*(2), 167–181. https://doi.org/10.1016/j.jecp.2008.04.002

Hartigan, P. M. (1985). Computation of the dip statistic to test for unimodality. *Applied Statistics*, *34*(3), 320–325. https://doi.org/10.2307/2347485

Hautala, J., Hyönä, J., & Aro, M. (2011). Dissociating spatial and letter-based word length effects observed in readers' eye movement patterns. *Vision Research*, *51* (15), 1719–1727. https://doi.org/10.1016/j.visres.2011. 05.015

Hautala, J., & Loberg, O. (2015). Breaking down the word length effect on readers' eye-movements. *Language, Cognition and Neuroscience*, *30*(8), 993–1007. https:// doi.org/10.1080/23273798.2015.1049187

Hermena, E. W., Liversedge, S. P., & Drieghe, D. (2017). The influence of a word's number of letters, spatial extent, and initial bigram characteristics on eye movement control during reading: Evidence from Arabic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(3), 451–471. https://doi.org/10.1037/ xlm0000319

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 166–190. https://doi.org/10.1037/a0033670

Huovilainen, T. M. (2018). Psycholinguistic descriptives. The Language Bank of Finland. http://urn.fi/urn:nbn:fi: lb-2018081601

Hyönä, J., & Bertram, R. (2011). Optimal viewing position effects in reading Finnish. *Vision Research*, *51*(11), 1279–1287. https://doi.org/10.1016/j.visres.2011.04. 004

Hyönä, J., Yan, M., & Vainio, S. (2018). Morphological structure influences the initial landing position in words during reading Finnish. *Quarterly Journal of Experimental Psychology*, *71*(1), 122–130. https://doi. org/10.1080/17470218.2016.1267233

Inhoff, A. W., Eiter, B. M., & Radach, R. (2005). Time course of linguistic information extraction from consecutive words during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 979–995. https://doi.org/10.1037/ 0096-1523.31.5.979

Inhoff, A. W., & Radach, R. (2014). Parafoveal preview benefits during silent and oral reading: Testing the parafoveal information extraction hypothesis. *Visual Cognition*, *22*(3–4), 354–376. https://doi.org/10.1080/ 13506285.2013.879630

Inhoff, A. W., Radach, R., Eiter, B. M., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *The Quarterly Journal of Experimental Psychology Section A*, *56*(5), 803–827. https://doi.org/ 10.1080/02724980244000639

Jacobs, A. M. (1987). On localization and saccade programming. *Vision Research*, *27*(11), 1953–1966. https:// doi.org/10.1016/0042-6989(87)90060-5

Khan, A., Loberg, O., & Hautala, J. (2017). On the eye movement control of changing reading direction for a single word: The case of reading numerals in Urdu. *Journal of Psycholinguistic Research*, *46*(5), 1273–1283. https://doi. org/10.1007/s10936-017-9491-1

Kliegl, R., & Engbert, R. (2005). Fixation durations before word skipping in reading. *Psychonomic Bulletin & Review*, *12*(1), 132–138. https://doi.org/10.3758/ BF03196358

Laurinavichyute, A. K., Sekerina, I. A., Alexeeva, S., Bagdasaryan, K., & Kliegl, R. (2019). Russian sentence corpus: Benchmark measures of eye movements in reading in Russian. *Behavior Research Methods*, *51*(3), 1161–1178. https://doi.org/10.3758/s13428-018-1051-6

Lenth, R, Singmann, H, Love, J, Buerkner, P, & Herve, M. (2018). Emmeans: Estimated marginal means, aka least-squares means. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Li, H. H., Barbot, A., & Carrasco, M. (2016). Saccade preparation reshapes sensory tuning. *Current Biology*, *26* (12), 1564–1570. https://doi.org/10.1016/j.cub.2016.04. 028

Li, X., & Pollatsek, A. (2020). An integrated model of word processing and eye-movement control during Chinese reading. *Psychological Review*, *127*(6), 1139–1162. https://doi.org/10.1037/rev0000248

Liu, Y., Guo, S., Yu, L., & Reichle, E. D. (2018). Word predictability affects saccade length in Chinese reading: An evaluation of the dynamic-adjustment model. *Psychonomic Bulletin & Review*, *25*(5), 1891–1899. https://doi.org/10.3758/s13423-017-1357-x

Liu, Y., Huang, R., Gao, D., & Reichle, E. D. (2017). Further tests of a dynamic-adjustment account of saccade targeting during the reading of Chinese. *Cognitive Science*, *41*, 1264–1287. https://doi.org/10.1111/cogs.12487

Liu, Y., Huang, R., Li, Y., & Gao, D. (2017). The word frequency effect on saccade targeting during Chinese reading: Evidence from a survival analysis of saccade length. *Frontiers in Psychology*, *8*(116), 1–8. https://doi. org/10.3389/fpsyt.2017.00116

Liu, Y., Yu, L., Fu, L., Li, W., Duan, Z., & Reichle, E. D. (2019). The effects of parafoveal word frequency and segmentation on saccade targeting during Chinese reading. *Psychonomic Bulletin & Review*, *26*(4), 1367–1376. https://doi.org/10.3758/s13423-019-01577-x

Liversedge, S. P., Drieghe, D., Li, X., Yan, G., Bai, X., & Hyönä, J. (2016). Universality in eye-movements and reading: A trilingual investigation. *Cognition*, *147*, 1–20. https:// doi.org/10.1016/j.cognition.2015.10.013

Ludden, T. M., Beal, S. L., & Sheiner, L. B. (1994). Comparison of the Akaike information criterion, the schwarz criterion and the F test as guides to model selection. *Journal of Pharmacokinetics and Biopharmaceutics*, *22*(5), 431–445. https://doi.org/10. 1007/BF02353864

Mächler, M. (2013). *Package 'diptest'. R package version 0.75–5. R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Marcet, A., & Perea, M. (2018). Visual letter similarity effects during sentence reading: Evidence from the boundary technique. *Acta Psychologica*, *190*, 142–149. https://doi. org/10.1016/j.actpsy.2018.08.007

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye-movement control during reading: I. The location of initial eye fixations on words. *Vision Research*, *28*(10), 1107–1118. https://doi.org/10.1016/ 0042-6989(88)90137-X

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye-movement control during

reading: II. Frequency of refixating a word. *Perception & Psychophysics*, *46*(3), 245–253. https://doi.org/10.3758/BF03208086

McConkie, G. W., & Yang, S. N. (2003). How cognition affects eye-movements during reading. In J. Hyönä, H. Deubel, & R. Radach (Eds.), *The mind's eye* (pp. 413–427). North-Holland.

McConkie, G. W., & Zola, D. (1984). Eye movement control during reading: The effect of word units. In W. Prinz, & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 63–74). Springer. https://doi.org/10.1007/978-3-642-69382-3_5

McDonald, S. A. (2006). Parafoveal preview benefit in reading is only obtained from the saccade goal. *Vision Research*, *46*(26), 4416–4424. https://doi.org/10.1016/j.visres.2006.08.027

McDonald, S. A., & Shillcock, R. C. (2005). The implications of foveal splitting for saccade planning in reading. *Vision Research*, *45*(6), 801–820. https://doi.org/10.1016/j.visres.2004.10.002

McGowan, V. A., & Reichle, E. D. (2018). The "risky" reading strategy revisited: New simulations using E-Z reader. *Quarterly Journal of Experimental Psychology*, *71*(1), 179–189. https://doi.org/10.1080/17470218.2017.1307424

McSorley, E., Gilchrist, I. D., & McCloy, R. (2019). The programming of sequences of saccades. *Experimental Brain Research*, *237*(4), 1009–1018. https://doi.org/10.1007/s00221-019-05481-7

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research*, *45*(17), 2201–2217. https://doi.org/10.1016/j.visres.2005.02.014

Plummer, P., & Rayner, K. (2012). Effects of parafoveal word length and orthographic features on initial fixation landing positions in reading. *Attention, Perception, & Psychophysics*, *74*(5), 950–963. https://doi.org/10.3758/s13414-012-0286-z

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*(1), 1–56. https://doi.org/10.1016/j.cogpsych.2005.06.001

Radach, R., Inhoff, A., & Heller, D. (2004). Orthographic regularity gradually modulates saccade amplitudes in reading. *European Journal of Cognitive Psychology*, *16*(1–2), 27–51. https://doi.org/10.1080/09541440340000222

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*(2), 211–236. https://doi.org/10.1016/0022-0965(86)90037-8

Rayner, K., Raney, G. E., & Sereno, S. C. (1996). Eye movement control in reading: A comparison of Two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(5), 1188–1200. https://doi.org/10.1037/0096-1523.22.5.1188

Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic Bulletin & Review*, *17*(6), 834–839. https://doi.org/10.3758/PBR.17.6.834

Rayner, K., Slattery, T. J., Drieghe, D., & Liversedge, S. P. (2011). Eye-movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 514–528. https://doi.org/10.1037/a0020990

R Core Team and Contributors Worldwide. (2002). *The R stats package*. R Foundation for Statistical Computing. Vienna. http://www.R-project.org.

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, *33*(2), 110–149. https://doi.org/10.1016/j.dr.2013.03.001

Reichle, E. D., Rayner, K., & Pollatsek, A. (1999). Eye movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research*, *39*(26), 4403–4411. https://doi.org/10.1016/S0042-6989(99)00152-2

Reichle, E. D., Rayner, K., & Pollatsek, A. (2012). Eye-movements in reading versus nonreading tasks: Using E-Z Reader to understand the role of word/stimulus familiarity. *Visual Cognition*, *20*(4-5), 360–390. https://doi.org/10.1080/13506285.2012.667006

Reichle, E. D., & Sheridan, H. (2015). E-Z reader: An overview of the model and two recent applications. In A. Pollatsek, & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 277–292). Oxford University Press.

Reilly, R. G., & O'Regan, J. K. (1998). Eye movement control during reading: A simulation of some word-targeting strategies. *Vision Research*, *38*(2), 303–317. https://doi.org/10.1016/S0042-6989(97)87710-3

Reilly, R. G., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, *7*(1), 34–55. https://doi.org/10.1016/j.cogsys.2005.07.006

Reingold, E. M., & Rayner, K. (2006). Examining the word identification stages hypothesized by the E-Z Reader model. *Psychological Science*, *17*(9), 742–746. https://doi.org/10.1111/j.1467-9280.2006.01775.x

Reingold, E. M., Sheridan, H., Meadmore, K. L., Drieghe, D., & Liversedge, S. P. (2016). Attention and eye-movement control in reading: The selective reading paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(12), 2003–2020. https://doi.org/10.1037/xhp0000291

Rice, M. L., Leske, D. A., Smestad, C. E., & Holmes, J. M. (2008). Results of ocular dominance testing depend on assessment method. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, *12*(4), 365–369. https://doi.org/10.1016/j.jaapos.2008.01.017

Risse, S., Hohenstein, S., Kliegl, R., & Engbert, R. (2014). A theoretical analysis of the perceptual span based on SWIFT simulations of the n+ 2 boundary paradigm. *Visual Cognition*, *22*(3–4), 283–308. https://doi.org/10.1080/13506285.2014.881444

Rolfs, M., & Carrasco, M. (2012). Rapid simultaneous enhancement of visual sensitivity and perceived contrast during saccade preparation. *Journal of Neuroscience*, *32*(40), 13744–13752. https://doi.org/10.1523/JNEUROSCI.2676-12.2012

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270–1281. https://doi.org/10.3758/BF03201199

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*(1), 5–35. https://doi.org/10.3758/s13414-011-0219-2

Schotter, E. R., Leininger, M., & von der Malsburg, T. (2018). When your mind skips what your eyes fixate: How forced fixations lead to comprehension illusions in reading. *Psychonomic Bulletin & Review*, *25*(5), 1884–1890. https://doi.org/10.3758/s13423-017-1356-y

Schotter, E. R., von der Malsburg, T., & Leininger, M. (2019). Forced fixations, trans-saccadic integration, and word recognition: Evidence for a hybrid mechanism of saccade triggering in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 677–688. https://doi.org/10.1037/xlm0000617

Singmann, H., Bolker, B., Westfall, J., Aust, F., Højsgaard, S., Fox, J., … Christensen, R. H. B. (2018). *Analysis of factorial experiments. R package. R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://CRAN.R-project.org/package=afex.

Snell, J., van Leipsig, S., Grainger, J., & Meeter, M. (2018). OB1-reader: a model of word recognition and eye movements in text reading. *Psychological Review*, *125*(6), 969–984. https://doi.org/10.1037/rev0000119

The Finnish N-gram Corpus, version 1 (FNC1). 2014. Distributed by the University of Helsinki on behalf of the FIN-CLARIN Consortium. http://www.helsinki.fi/finclarin/fnc1.

Veldre, A., & Andrews, S. (2014). Lexical quality and eye movements: Individual differences in the perceptual span of skilled adult readers. *The Quarterly Journal of Experimental Psychology*, *67*(4), 703–727. https://doi.org/10.1080/17470218.2013.826258

Veldre, A., Reichle, E. D., Wong, R., & Andrews, S. (2020). The effect of contextual plausibility on word skipping during reading. *Cognition*, *197*, 104184. https://doi.org/10.1016/j.cognition.2020.104184

Vitu, F. (2003). The basic assumptions of E-Z Reader are not well-founded. *Behavioral and Brain Sciences*, *26*(4), 506–507. https://doi.org/10.1017/S0140525X0351010X

Warrington, K. L., McGowan, V. A., Paterson, K. B., & White, S. J. (2018). Effects of aging, word frequency, and text stimulus quality on reading across the adult lifespan: Evidence from eye-movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(11), 1714–1729. https://doi.org/10.1037/xlm0000543

Wei, W., Li, X., & Pollatsek, A. (2013). Word properties of a fixated region affect outgoing saccade length in Chinese reading. *Vision Research*, *80*, 1–6. https://doi.org/10.1016/j.visres.2012.11.015

White, S. J. (2007). Foveal load and parafoveal processing: The case of word skipping. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye-movements: A window on mind and brain* (pp. 409–424). Elsevier.

White, S. J., & Liversedge, S. P. (2006). Foveal processing difficulty does not modulate non-foveal orthographic influences on fixation positions. *Vision Research*, *46*(3), 426–437. https://doi.org/10.1016/j.visres.2005.07.006

White, S. J., & Staub, A. (2012). The distribution of fixation durations during reading: Effects of stimulus quality. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 603–617. https://doi.org/10.1037/a0025338

Wotschack, C., & Kliegl, R. (2013). Reading strategy modulates parafoveal-on-foveal effects in sentence reading. *The Quarterly Journal of Experimental Psychology*, *66*(3), 548–562. https://doi.org/10.1080/17470218.2011.625094

Yang, S. N., & Vitu, F. (2007). Dynamic coding of saccade length in reading. In R. P. G. Van Gompel, M. F. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements* (pp. 293–317). Elsevier. https://doi.org/10.1016/B978-008044980-7/50015-X.

Yao-N'Dré, M., Castet, E., & Vitu, F. (2014). Inter-word eye behaviour during reading is not invariant to character size: Evidence against systematic saccadic range error in reading. *Visual Cognition*, *22*(3-4), 415–440. https://doi.org/10.1080/13506285.2014.886652

Yu, L., Liu, Y., & Reichle, E. D. (2020). A corpus-based vs. Experimental examination of word- and character-frequency effects in Chinese reading: Theoretical implications for models of reading. *Journal of Experimental Psychology: General*. Advance online publication.