Accurate line-initial fixations but not line-final fixations differ from intra-line fixations during both reading and z-string scanning

Adam J. Parker[1], Muchan Tao[2], Martin R. Vasilev[1]

Affiliations: [1]Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, London, WC1H 0AP; [2]Department of Language and Cognition, Division of Psychology and Language Sciences, University College London, Chandler House, 2 Wakefield Street, WC1N 1PF.

Correspondence: Correspondence regarding this article should be addressed to Adam J. Parker, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, 26 Bedford Way, London, WC1H 0AP. Email: adam_parker@ucl.ac.uk.

Data availability statement: The materials and the data sets generated and analysed are available in the Open Science Framework (OSF) repository, https://osf.io/tpf8e/. This repository also includes an R Markdown script to reproduce all analyses and generate the manuscript.

Ethical approval statement: The experimental procedure was granted ethical approval by the UCL Department of Experimental Psychology's Ethics Chair, ethics application number: EP/2021/015.

ORCID iD: Adam J. Parker: 0000-0002-1367-2282, Martin R. Vasilev: 0000-0003-1944-8828

Abstract

Return-sweeps, which move the reader's gaze from the end of one line to the beginning of the next, typically result in shorter line-final fixations and longer accurate line-initial fixations compared to intra-line fixations. The mechanisms underlying these differences have been widely debated. To assess linguistic and oculomotor contributions to these return-sweep fixation differences, we compared the eye movements of 41 participants during normal reading and z-string scanning, an oculomotor control condition to reading, which is devoid of useful linguistic content. Our results indicate that line-final fixations are shorter and accurate line-initial fixations are longer under both tasks, underscoring the significant role of the oculomotor system in determining fixation durations across tasks. Notably, the reduction in line-final fixation durations compared to intra-line fixations did not differ between tasks. This suggests that oculomotor coordination or visual processing, rather than linguistic processing, drives shorter line-final fixations. In contrast, the difference in accurate line-initial fixation durations between reading and z-string scanning implies that longer accurate line-initial fixations are likely a result of lexical processing and oculomotor coordination or visual processing. These findings advance our understanding of eye movement control by highlighting the combined influence of linguistic and oculomotor processes on return-sweep fixation durations.

Keywords: eye movements, reading, z-string scanning, return-sweeps, oculomotor coordination

In recent years, research on return-sweep saccades during reading has significantly increased (Adedeji et al., 2022; Christofalos et al., 2024; Parker et al., 2017; Parker, Slattery, et al., 2019; Parker, Nikolova, et al., 2019; Parker & Slattery, 2019, 2021, 2024; Parker et al., 2020; Parker et al., 2023; Slattery & Parker, 2019; Slattery & Vasilev, 2019; Vasilev et al., 2021; Wang et al., 2024). Return-sweeps are large saccadic eye movements that move readers' gaze from one line to the next, significantly impacting fixation durations before and after the return-sweep. Consistent findings show that line-final fixations (fixations just before a return-sweep) are shorter than intra-line fixations (fixations within a line), while accurate line-initial fixations (fixations at the start of a line followed by a rightwards pass) are longer than intra-line fixations (Abrams & Zuber, 1972; Hawley et al., 1974; Heller, 1982; Hofmeister, 1998; Rayner, 1977, 1978). Several theories have been proposed to explain these fixation duration differences, with explanations generally clustering on lexical processing or oculomotor/visual accounts. To date, no study has examined how lexical processing and oculomotor coordination/visual processing contribute to shorter line-final fixations and longer accurate line-initial fixations (relative to intra-line fixations). To this end, we compared participants' eye movements as they read multi-line texts and scanned rows of letter strings under a *z-reading paradigm* (Vitu et al., 1995), an oculomotor control condition devoid of useful linguistic content. If differences in fixation duration across fixation types were the same across both tasks then this would indicate return-sweep fixation duration differences are the consequence of oculomotor coordination or visual processing rather than lexical processing.

Return-sweep saccades typically launch from 4-8 characters from the line's end in alphabetic reading (Hofmeister et al., 1999; Parker & Slattery, 2019; Parker et al., 2020; Rayner, 1998). Numerous studies have confirmed that line-final fixations are shorter than intra-line fixations (Abrams & Zuber, 1972; Adedeji et al., 2022; Christofalos et al., 2024; Hawley et al., 1974; Heller, 1982; Parker, Slattery, et al., 2019; Parker, Nikolova, et al., 2019; Parker & Slattery, 2021; Parker et al., 2023; Rayner, 1977; Slattery & Parker, 2019; Vasilev et al., 2021; Wang et al., 2024). Mitchell et al. (2008) suggested that these shorter fixations result from return-sweep preparation, with the primary purpose of line-final fixations being to orient the visual system. An extreme version of this theory posits that line-final fixations do not involve linguistic processing. Supporting this, Hofmeister (1998) found that text degradation did not affect line-final fixation durations. However, ample evidence now suggests otherwise, as lexical frequency has been shown to influence line-final fixations and reading times on line-final words (Parker et al., 2023; Parker & Slattery, 2024).

Alternative explanations for shorter line-final fixations have been proposed. Rayner (1977) suggested that the absence of a word to the right of fixation eliminates the need to process parafoveal information, thus shortening line-final fixations. Similarly, shorter line-final fixations could reflect a reduction in skipping costs or reduced lateral masking at the end of a line. Alternatively, readers may terminate line-final fixations earlier as they can conduct additional lexical processing during the return-sweep, which is longer than typical intra-line reading saccades (see Parker & Slattery, 2024, for a discussion).

Return-sweeps, like any saccade, are prone to systematic and random error (McConkie et al., 1998). They undershoot their target 40-60% of the time, necessitating an immediate corrective saccade towards the left margin (Slattery & Vasilev, 2019). Consequently, return-sweeps have two possible outcomes: accurate line-initial fixations, where the return-sweep is followed by a rightwards pass, or under-sweep fixations, where readers land short of their intended target and make a leftwards corrective saccade before a rightwards pass.

Accurate line-initial fixations, which land 4-8 characters from the start of the line, are longer than intra-line fixations (Adedeji et al., 2022; Christofalos et al., 2024; Parker, Slattery, et al., 2019; Parker, Nikolova, et al., 2019; Parker & Slattery, 2021; Parker et al., 2020; Parker et al., 2023; Slattery & Parker, 2019; Wang et al., 2024). Several theories have been proposed to explain this. Parker et al. (2017) suggested that the absence of parafoveal preview for line-initial words, which lie outside the perceptual span prior to fixation, might result in longer line-initial fixations. Alternatively, Rayner (1978) and Kuperman et al. (2010) suggested that longer accurate line-initial fixations might result from establishing a mode of saccadic programming after the return-sweep.

Under-sweep fixations are typically shorter than intra-line fixations and are generally assumed to involve little lexical processing (Hawley et al., 1974; Hofmeister, 1998; Shebilske, 1975). These fixations are thought to be primarily due to oculomotor error, with the main goal being to rapidly plan and execute a corrective saccade to the intended target of the return-sweep (Becker, 1976). While studies have reported that under-sweep fixation durations are not influenced by the properties of the fixated word (Parker et al., 2020; Slattery & Parker, 2019), there is evidence that readers utilise this pause before a corrective saccade to extract information from the undershot line-initial word and fixated word that facilitates subsequent processing (Parker & Slattery, 2019; Parker et al., 2020; Slattery & Parker, 2019). Previously, in a study comparing reading and letter scanning, Hofmeister (1998) reported that under-sweep durations did not differ between tasks, suggesting that the impact of lexical processing during an under-sweep fixation may be minimal.

The *z-reading paradigm* offers a way to assess linguistic and oculomotor/visual contributions to return-sweep fixation differences. In this paradigm, participants read strings of meaningless letters resembling real text (e.g., *Eye movements during reading* -> *Xxx xxxxxxxxx xxxxxx xxxxxxx*), preserving the text's spatial layout but removing higher-level linguistic information. This provides an excellent oculomotor control condition for reading. The *z-reading paradigm* has been associated with longer fixation durations (Al-Zanoon et al., 2017; Gagl et al., 2022; Rayner & Fischer, 1996; Vitu et al., 1995), increased skipping for longer letter strings

(Rayner & Fischer, 1996; Vitu et al., 1995), and fewer regressions (Nuthmann et al., 2007). The paradigm has previously been used to examine whether shorter under-sweep fixations are the result of general oculomotor coordination processes (Hofmeister, 1998).

Hofmeister (1998) compared eye movements during reading and z-string scanning. Their analyses were primarily concerned with landing positions, where it was reported that initial landing positions of line-initial fixations were further from the margin during scanning than reading. However, they also compared under-sweep fixation durations, noting no differences between tasks, enabling Hofmeister to conclude that under-sweep fixations are almost exclusively governed by oculomotor control. Note, however, that Hofmeister did not compare line-final or accurate line-initial fixations between tasks. We, therefore, aimed to use the *z-reading paradigm* to differentiate between linguistic and oculomotor/visual contributions to return-sweep fixation duration differences.

Pre-Registered Research Questions and Predictions

We pre-registered the following predictions:

Return-sweep fixation types during paragraph reading

Within our statistical modelling framework, we applied a coding scheme that enabled us to first compare return-sweep fixations with intra-line reading fixations during reading. Our questions and predictions are as follows:

- *Are line-final reading fixations shorter than intra-line reading fixations?* We predicted shorter line-final reading fixations relative to intra-line reading fixations.
- *Are accurate line-initial reading fixations shorter than intra-line reading fixations?* We predicted longer accurate line-initial reading fixations relative to intra-line reading fixations.
- *Are under-sweep reading fixations shorter than intra-line reading fixations?* We predicted shorter under-sweep reading fixations relative to intra-line reading fixations.

Differences between fixation types between z-string scanning and multiline text reading

Within our statistical models, task is coded to compare fixations during z-string scanning to multiline text reading. As such, our predictions are qualified by interactions within statistical models. Our questions and predictions are as follows:

- *Do intra-line reading fixation durations differ from z-string scanning fixation durations?* Previous studies have reported longer fixations during z-string scanning than during reading (e.g., Rayner & Fischer, 1996). Therefore, we predicted longer intra-line fixations during z-string scanning (i.e., a significant simple effect of task).

- *Does the reduction in duration for line-final fixations (relative to intra-line fixations) differ between reading and z-string scanning?* If shorter fixations during reading result from lexical processing, then we anticipate similar durations between intra-line fixations and line-final fixations during scanning, which results in an interaction between fixation type and task. If, however, shorter line-final fixations during reading are driven by oculomotor coordination/visual processing then we would expect shorter line-final fixations across both tasks and no interaction when comparing data across tasks.

- *Does the increase in duration for accurate line-initial fixations (relative to intra-line fixations) differ between reading and z-string scanning?* If longer accurate line-initial fixations during reading result from linguistic processing we would expect similar durations between intra-line and accurate line-initial fixations during scanning, resulting in an interaction between fixation type and task. This is because readers will be able to engage in lexical processing at the start of a new line during reading but not scanning. If, however, longer accurate line-initial fixations during reading are driven by oculomotor coordination/visual processing then we would expect longer accurate line-initial fixations across both tasks and no interaction when comparing data across tasks.

- *Does the reduction in duration for under-sweep fixations (relative to intra-line fixations) differ between reading and z-string scanning?* If readers engage in lexical processing during an under-sweep fixation that facilitates their subsequent reading behaviour (Parker & Slattery, 2019; Parker et al., 2020; Slattery & Parker, 2019), then we might observe a slight difference in the reduction in durations for under-sweep fixations relative to intra-line fixations across tasks (i.e., a significant interaction). However, given that under-sweep fixations are generally considered to be under oculomotor control (Hofmeister, 1998), we may alternatively observe similar reductions for under-sweep fixations (relative to intra-line fixations) across both tasks and, therefore, a lack of interaction with task.

Methods

This experiment was pre-registered on the Open Science Framework (OSF) before data collection. The registration form, task materials, analysis scripts, and anonymised data are available on the Open Science

Framework: https://osf.io/tpf8e/.

Participants

A priori power analyses were conducted for all fixed effects of interest for our comparison of reading and scanning fixations within a frequentist linear mixed modelling framework. We started by simulating multi-level data for 40 statistical subjects, where each statistical subject had data for 30 trials of text reading and 30 trials of z-string scanning and characters were displayed across four lines in each trial. The fixation durations for each fixation type during text reading were taken from Parker and Slattery's (2021) short line condition as the line lengths were comparable: intra-line fixations: 200.6 ms, line-final fixations: 191.4 ms, accurate line-initial fixations: 257.9 ms, and under-sweep fixations: 148.9 ms. For z-string scanning, we simulated a 38 ms increase in fixation duration for intra-line reading (Rayner & Fisher, 1996) such that the total duration equated to approximately 238.6 ms. Under the linguistic account, we would expect that return-sweep fixations should not differ from intra-line fixations during scanning. Hence, we simulated data where there is a negligible effect of fixation type for z-string scanning. We simulated this data 1,000 times and, on each run, fitted a linear mixed-effects model to the data *(log10(fixation duration)~ fixation type* $\times$ *stimuli type + (1 | participant) + (1 | item))*, tallying each time a significant result was obtained for each fixed effect. Simulations suggested that 40 participants would provide sufficient power to detect all critical interactions where we predicted a difference with sufficient power (i.e., >90%) using a significance threshold of $|t|>2$.

To reach our pre-registered sample size, we initially recruited 55 participants via the UCL Psychology and Language Sciences SONA Participant Pool. Participants were aged between 18 and 45 years old, had spoken English for a minimum of 10 years, had no language, hearing, or visual impairments, and had no history of neurological illness. Participants were reimbursed at a rate of £9.00/hour or received course credit for their participation. We imposed several data cleaning procedures that resulted in a final sample of 41 participants. For more information on the data cleaning procedures, see *Data Cleaning and Final Sample.*

The experimental procedure was granted ethical approval by the UCL Department of Experimental Psychology's Ethics Chair, ethics application number: EP_2021_015.

Reading Task

Thirty passages were taken from the Provo Corpus (Luke & Christianson, 2018). On average, the paragraphs were 49.97 words long (SD_{words}= 5.80; $range_{words}$: 40–59). The mean word length was 4.75 letters ($SD_{letters}$= 2.51; $range_{letters}$: 1-15). Words in each passage had an average Zipf frequency of 5.71 (SD_{zipf}= 1.42; $range_{zipf}$= 1.17-7.67) based on the SUBTLEX-UK Corpus (van Heuven et al., 2014) and an average cloze probability of

0.20 (SD_{cloze}= 0.20; $range_{cloze}$= 0.00-1.00). Each paragraph was 2.63 sentences long on average ($SD_{sentences}$= 0.96; $range_{sentences}$: 1-5 sentences) and was displayed across 4.13 lines on average (SD_{lines}= 0.51; $range_{lines}$: 3-5). During the passage reading task, participants were instructed to read silently for comprehension while their eye movements were recorded. After reading each paragraph, participants were asked a single comprehension question with three options (see Figure 1).

Text Reading

Very similar, but even more striking, is the evidence from athletic

training. As with rehearsing a piece on the piano, practicing a

complex physical task in the mind alone is nearly as effective a

learning strategy as actually physically doing it. But it does

not stop there.

Q: Where is the striking evidence from?
1. Music rehearsal
2. Athletic training
3. School learning

Z-String Scanning

zzzz zzzzzzzz zzz zzzz zzzz xzzzzxzzz zz zzz zzzzzzzz zzzz zzzzzzzz

xzzzzzzzz zz zzzz zzzzzzzzzz z zzzzz zz zzz zzzzzz zzzzzzzzzz z

zxzzzzz zzzzzzzz zzzz zz zzz zzzz zzzzz zz zzzzzz xz zzzzzzzzz z

zzxzzzzz zzzzzzzz zz zzzzzzzz zzzxzzzzzz zzzzz zzz zzz xz zzzz

zzz zzzz xzxxzz

Q: How many ‘x’s are in the string?
1. 11
2. 7
3. 9

Figure 1: Example stimuli and question for text reading and z-string scanning.

Scanning Task

This scanning task is based on the *z reading paradigm* (Vitu et al., 1995). The characters in the paragraphs from the Reading Task were replaced with the letter *z*, preserving capitalization and empty spaces (but not punctuation, to avoid distractors in the search task), such that words and z-strings were matched on length. The letter *x* was randomly inserted 5-15 times in the string of *z*s. Participants were instructed to scan the string of letters (from left to right) and count how many times the letter *x* appears in it. After each trial, participants had to select the correct number of *x*s from three options.

Apparatus

Eye movements were recorded using an SR Research EyeLink 1000-Plus eye-tracker, which sampled at 1000 Hz. While viewing was binocular, only the right eye was tracked. To minimise head movements, a chin-and-forehead rest was used. Stimuli were presented on a 23.8" Dell G2422HS LCD monitor (resolution: 1920×1080) in 18 pt. Courier New font as black text over a white background. The eye-to-screen distance was 84 cm such that each letter subtended 0.26° horizontally. The experiment was programmed in SR Research Experiment Builder and was run on a Windows 11 PC.

Procedure

The experiment started with a 9-point calibration and validation procedure. Calibration accuracy was kept at $<0.4^\circ$ across the experiment. Drift checks were presented before every trial and participants were recalibrated whenever necessary, but at least every 15 trials. Participants were randomly allocated to complete the reading task first followed by the scanning task or vice versa. For the scanning task participants were explicitly instructed to adopt the same left-to-right strategy that is typical for English reading. Each task started with two practice trials, followed by 30 experimental trials. Participants were offered breaks whenever recalibration occurred. Each trial started with a fixation point that appeared to the left of the first character on the first line. Once a stable fixation was detected, the experimenter started the trial. When a participant had finished reading or scanning they pressed the *space bar* to terminate the trial and then answered the multiple choice question by pressing either *1*, *2*, or *3* on the keyboard.

Data Analysis

Data Cleaning and Final Sample

We pre-registered that participants needed to score 70% or more on the reading comprehension questions. This led to the removal of six participants. We additionally removed one participant's data as they failed to complete the study and a further seven due to calibration issues, excessive blinking, poor quality data, or corrupted files. The final sample consisted of 41 participants (31 female) with a mean age of 22.39 years (SD_{years}= 3.38).

The data of the remaining 41 participants were pre-processed using the popEye package (version 0.8.1; Schroeder, 2019) within R (version 4.4.1; R Development Core Team, 2020). Fixations were automatically vertically aligned against the text using the *chain* method (Carr et al., 2022). Fixations less than 50 ms were combined with the next fixation if they were within 1 character from each other. We pre-registered that we would remove trials in which participants made five or more blinks, leading to the removal of 20.81% of trials. For the remaining trials, fixations preceded or followed by a blink were removed as were fixations that were shorter than 50 ms or longer than 1200 ms, resulting in the removal of 4.96% of fixations. We then applied a Hoaglin and Iglewicz (1987) outlier removal procedure to reading time data to identify outliers individually for each participant across each statistical condition. This procedure defined outliers as data points that were 2.2 times the difference between the first quartile (*Q1*) and the third quartile (*Q3*), above or below the *Q1* and *Q3* values (e.g., lower boundary = *Q1* − 2.2 × (*Q3*-*Q1*); upper boundary = *Q3* + 2.2 × (*Q3*-*Q1*)). This led to the removal of 1.88% of fixations.

Registered Confirmatory Analysis of Return-Sweep Fixation Duration

For our pre-registered analyses, a series of linear mixed-effects models were fitted to *log10* transformed data using the *lmer()* function from the lme4 package (version 1.1.35.3; Bates et al., 2015). The model comparing fixation durations between task type adopted the structure *dv~ Task × Fixation Type + (1 + Task × Fixation Type | participant) + (1 + Task × Fixation Type | item)*, where participant and item are random factors. Treatment coding was utilised so that data for intra-line fixations from the reading task represented the intercept to which return-sweep fixations across the two tasks were compared. To specifically examine return-sweep fixation durations during scanning, we fitted an additional exploratory model to scanning data: *dv~ Fixation Type + (1 + Fixation Type | participant) + (1 + Fixation Type | item)*. For all models, we report regression coefficients (*b*), standard errors (*SE*), and *t*-values.

To estimate the best-fitting random structure for each model, the *buildmer()* function from the buildmer package (version 2.11; Voeten, 2021) was used. First, a maximal structure was fitted to the data before applying a backwards elimination process based on the significance of the change in log likelihood between models. The most basic and possible model retained all fixed effects and random intercepts for participants and items.

To evaluate the evidence for the critical null effects, we supplemented our analyses with Bayes Factor analysis. Bayes Factors quantify how much evidence the data (and priors) provide in favour of two competing models and allow us to infer how much a given hypothesis is consistent with the data (for reviews see Nicenboim et al., 2023, and Wagenmakers, 2017). Bayes Factors were computed by first fitting Bayesian linear-mixed effects models to fixation duration data using the *brm()* function from the brms package (version 2.21.0; Bürkner, 2007). The models included the same fixed effects as the *lmer()* models. Non-informative priors *normal(0,1)* were assumed for each fixed effect. Each model used 12,000 iterations with four chains, where the first 2,000 iterations were discarded due to warm-up. Then the *hypothesis()* function was implemented to calculate the Bayes Factors (BF_{10}) for each fixed effect. The *hypothesis()* function computes Bayes Factors using the Savage-Dickey density ratio method (Dickey, 1971), where Bayes Factors for individual parameters within a model are taken as the posterior density of the model parameter of interest divided by the prior density at the critical point of inference (e.g., zero if assessing whether an estimate is not equal to zero).

The combination of frequentist and Bayesian analysis enabled us to take a two-stage approach to inference. We considered results to be statistically significant where $|t| > 2$. If $|t| < 2$ and $BF_{10} > 1/3$, we considered there to be insufficient evidence. If $|t| < 2$ and $BF_{10} < 1/3$, we concluded that there was evidence in favour of the null hypothesis.

Non-Registered Exploratory Analysis of Return-Sweep and Corrective Saccade Parameters

For completeness, we analysed several return-sweep and corrective saccade parameters: *return-sweep launch position* (character position relative to the end of the line), *probability of making an under-sweep fixation*, *landing position of accurate line-initial fixations* (character position relative to the start of the line), and *landing position of under-sweep fixations* (character position relative to the start of the line). For the analyses of launch position and landing positions, we removed data points where the fixation was either more than 30 characters from the end of a line or more than 30 characters from the start of the line. For each measure, we fitted (generalised) linear mixed-effects models using the *(g)lmer()* function from the lme4 package. The model comparing parameters between reading and scanning was specified as *dv~ Task + (1 + Task | participant) + (1 + Task | item)*, where participants and items are random factors. Treatment coding was utilised so that data from the reading task represented the intercept to which scanning data were compared. As with our registered analyses, we used the *buildmer()* function to determine the random effects structure and combined frequentist and Bayes Factor analysis to adopt a two-stage approach to inference.

Results

Task Accuracy

Participants' task accuracy was lower during reading 84.57% (*SD*= 36.13%) than during scanning 95.7% (*SD*= 20.39%), b= 1.23, *SE*= 0.39, z= 3.19, BF_{10}= 34.63.

Eye Movements

Registered Confirmatory Analysis of Return-Sweep Fixation Duration

We report a confirmatory analysis of fixation durations where we compared return-sweep fixations to intra-line reading fixations across tasks. Mean fixation durations are reported in Table 1 and distributions are visualised in Figure 2.

Table 1: Mean Return-Sweep Fixation Durations per Task.

Fixation Type	Reading	Scanning
Intra-Line	216 (76)	248 (88)
Line-Final	186 (84)	216 (96)
Accurate Line-Initial	241 (81)	260 (92)
Under-Sweep	164 (48)	184 (63)

Note: [a]Standard deviations are shown in parentheses.

The model fitted to log-transformed fixation duration data *(lmer(dv~ Task × Fixation Type + (1 | Participant) + (1 + Task | Item)))* indicated that line-final reading fixations and under-sweep reading fixations were shorter than intra-line reading fixations, while accurate line-initial reading fixations were longer than intra-line reading fixations (see Table 2). The simple effect of Task indicated that intra-line fixations were longer during scanning than reading. The difference between intra-line and line-final fixation durations did not differ between reading and scanning. However, interactions in the model indicate that the increase in duration of accurate line-initial fixations compared to intra-line fixations was smaller during scanning than reading and

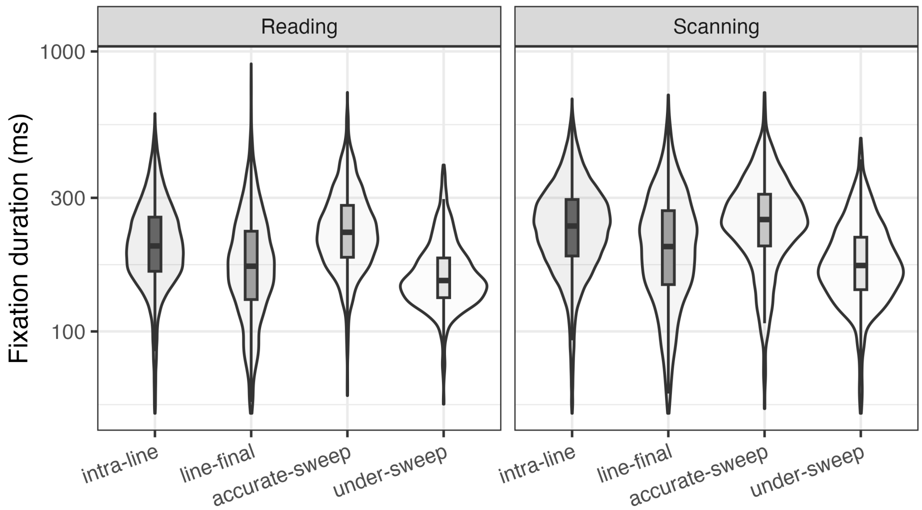

Figure 2: Return-sweep fixation durations per task. The box extends from the first to the third quartile with the line in the middle representing the median.

293 that the decrease in duration of under-sweep compared intra-line fixations was greater for scanning than
294 reading.

Table 2: Linear Mixed-Effects Results and Bayes Factors for Return-Sweep Fixation Durations.

Dataset	Fixed Effect	b	SE	t	BF10
Reading and Scanning	(Intercept)	2.30	0.01	315.31	-
	Fixation Type [Line-Final]	0.06	<0.01	58.22	6.64e+14
	Fixation Type [Accurate]	-0.08	<0.01	-33.31	5.05e+19
	Fixation Type [Under-Sweep]	0.05	<0.01	17.95	1.09e+15
	Task [Scanning]	-0.11	<0.01	-31.11	2.20e+24
	Fixation Type [Line-Final] × Task [Scanning]	0.01	<0.01	1.70	1.32e-02
	Fixation Type [Accurate] × Task [Scanning]	-0.03	<0.01	-6.81	6.60e+14
	Fixation Type [Under-Sweep] × Task [Scanning]	-0.01	<0.01	-2.50	1.16e-01
Scanning	(Intercept)	2.36	0.01	298.38	-
	Fixation Type [Line-Final]	-0.07	<0.01	-29.05	3.43e+15
	Fixation Type [Accurate]	0.03	<0.01	9.36	2.03e+14
	Fixation Type [Under-Sweep]	-0.12	<0.01	-35.37	3.43e+15

To examine whether the difference between intra-line and accurate line-initial fixations was small but reliable or completely abolished for scanning, we fitted a supplemental model to scanning data *(lmer(dv~ Fixation Type + (1 | Participant) + (1 | Item)))*. The model indicated that line-final and under-sweep fixations were shorter than intra-line fixations, and accurate line-initial fixations were longer than intra-line reading fixations.

Non-Registered Exploratory Analysis of Return-Sweep and Corrective Saccade Parameters

We report exploratory, non-registered analyses for four return-sweep and corrective saccade parameters. Descriptive statistics are reported in Table 3 and distributions are visualised in Figure 3.

Table 3: Mean Return-Sweep and Corrective Saccade Parameters per Task.

Parameter	Reading	Scanning
Launch Position	8.79 (5.33)	7.80 (6.14)
p(Under-Sweep Fixation)	41.98 (49.36)	45.67 (49.82)
Accurate Landing Position	6.61 (4.93)	5.56 (4.46)
Under-Sweep Landing Position	8.84 (4.30)	8.00 (4.00)

Note: [a]Standard deviations are shown in parentheses.

First, for return-sweep launch position, a linear mixed-effects model *(lmer(dv~ Task + (1 + Task | Participant) + (1 | Item)))* indicated no significant difference between tasks (see Table 4). The Bayes Factor for the fixed effect of Task indicated that there was insufficient evidence to draw a decisive conclusion. Second, for the probability of making an under-sweep fixation, the generalised linear mixed-effects model *(glmer(dv~ Task + (1 + Task | Participant) + (1 + Task | Item)))* indicated no significant difference between tasks. Again the Bayes Factor indicated that there was insufficient evidence to draw a decisive conclusion. Third, for accurate line-initial landing position, the linear mixed-effects model *(lmer(dv~ Task + (1 + Task | Participant) + (1 | Item)))* indicated that readers' accurate line-initial fixations landed closer to the left margin for scanning than text reading. Finally, under-sweep landing position, the linear mixed-effects model *(lmer(dv~ Task + (1 + Task | Participant) + (1 | Item)))* indicated no significant difference between tasks. The Bayes Factor indicated insufficient evidence to draw a decisive conclusion.

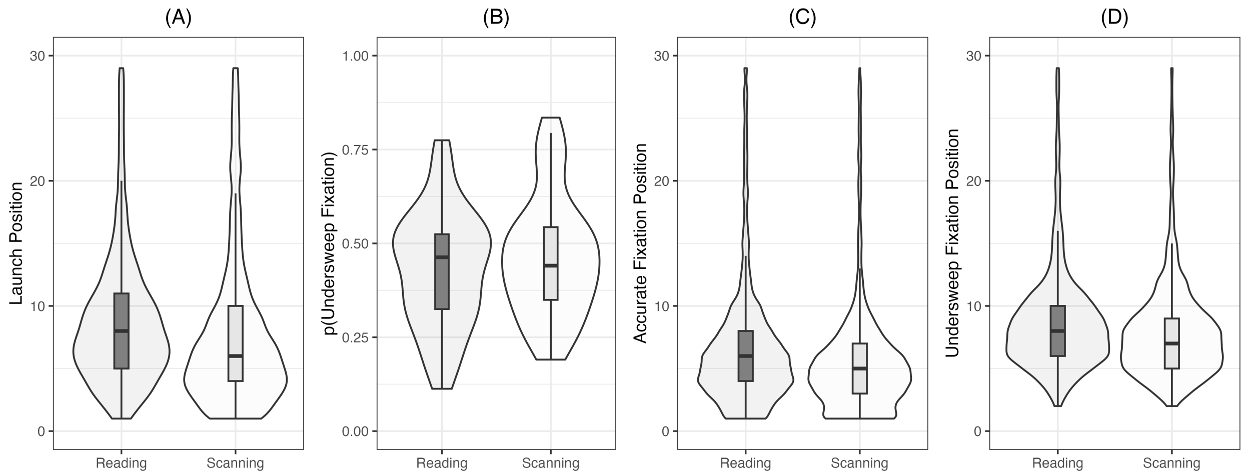

Figure 3: Return-sweep parameters: (A) Return-sweep launch position (characters from the end of a line); (B) Proportion of under-sweep fixations; (C) Accurate return-sweep landing position (characters from the start of the line); and (D) Under-sweep landing position (characters from the start of the line). The box extends from the first to the third quartile with the line in the middle representing the median.

Table 4: Linear Mixed-Effects Results and Bayes Factors for Return-Sweep and Corrective Saccade Parameters.

Measure	Fixed Effect	b	SE	t/z	BF10
Launch Position	(Intercept)	8.40	0.36	23.17	-
	Task	-0.56	0.39	-1.43	0.87
p(Under-Sweep Fixation)	(Intercept)	-0.33	0.11	-2.96	-
	Task	0.18	0.12	1.48	0.33
Accurate Landing Position	(Intercept)	6.80	0.29	23.50	-
	Task	-1.27	0.33	-3.79	72.99
Under-Sweep Landing Position	(Intercept)	8.63	0.31	27.60	-
	Task	-0.50	0.34	-1.45	1.03

Note: [a]The Bayes Factor for p(Under-Sweep Fixation) was rounded down from 0.335.

Discussion

To examine linguistic and oculomotor/visual processing contributions to return-sweep fixation durations, 41 participants read 30 passages of text for comprehension and scanned 30 z-letter strings for xs. First, we replicated the well-established finding that relative to intra-line reading fixations, line-final fixations and under-sweep reading fixations are shorter in duration while accurate line-initial reading fixations are longer (e.g., Abrams & Zuber, 1972; Adedeji et al., 2022). Second, we compared fixation duration differences between intra-line fixations and return-sweep fixations across reading and scanning. The novel contributions of our work can be summarised in three general points. First, the reduction in line-final fixation durations, relative to intra-line fixations, did not differ across tasks. Second, the increase in fixation duration for accurate line-initial fixations, relative to intra-line fixations, was smaller during scanning than reading. Third, the reduction in under-sweep fixation durations, relative to intra-line fixations, was larger during scanning than reading. We discuss each point in turn.

Our mixed-effects analysis of data from both tasks replicates a widely reported finding in the literature that intra-line fixations are longer during scanning than reading (e.g., Al-Zanoon et al., 2017; Rayner & Fischer, 1996). However, what is also clear from the data is that this increase in fixation duration extends to return-sweep fixations, where line-final, accurate line-initial and under-sweep fixations were longer during scanning than reading.

When comparing the difference between intra-line reading fixations and line-final fixations across tasks, the reduction in duration did not differ between reading and scanning. Given that there was no meaningful linguistic content during scanning, we can rule out the suggestion that reduced lexical processing or additional time to conduct lexical processing during the return-sweep drives shorter line-final fixations. Instead, this reduction is likely driven by oculomotor or visual processing. Given that both tasks have similar oculomotor and visual processing demands, we cannot pinpoint the exact cause of shorter line-final fixations. Previously, Hofmeister (1998) suggested that the primary purpose of a line-final fixation is to programme the return-sweep. This may indeed be the case, but given previous evidence (Parker et al., 2023; Parker & Slattery, 2024) it should be made clear that this does not come at a cost to lexical processing. Future research will need to tease this account apart from those which claim the reduction stems from a lack of information to the right of fixation (i.e., lateral masking or reduced parafoveal processing).

Our results indicated that the increase in duration for accurate line-initial fixations, relative to intra-line fixations, was smaller during scanning than during reading. A supplemental model fitted to the scanning data confirmed that, while this increase was smaller for scanning, accurate line-initial fixations were still longer than intra-line fixations. If longer accurate line-initial fixations during reading were driven purely by oculomotor/visual processing, we might have expected the same increase for accurate line-initial fixations across both tasks. By contrast, if this increase stemmed from the processing of meaningful linguistic content, then we may have expected a *Fixation Type* × *Task* interaction in our pre-registered model and no difference in fixation duration between accurate line-initial fixations and intra-line reading fixations in our supplemental model fitted to scanning data. The data supported neither of these predictions. The most parsimonious explanation here is then that both linguistic and oculomotor/visual processing contribute to longer accurate line-initial fixations. Perhaps this reflects a combination of saccade planning and delayed lexical access driven by the lack of parafoveal preview prior to direct fixation.

Our comparative analysis of data from both tasks also indicated that the reduction in under-sweep fixation durations, relative to intra-line fixations, was larger during scanning than reading. Recent evidence indicates that lexical processing for line-initial words can occur during an under-sweep and that readers can acquire useful information that informs subsequent reading of words receiving an under-sweep fixation (Parker & Slattery, 2019; Parker et al., 2020; Slattery & Parker, 2019). One plausible outcome was that the reduction in duration for under-sweep reading fixations, relative to intra-line reading fixations, may have been smaller during scanning given that there was no meaningful linguistic content. However, the data were in the opposite direction. So why might the reduction in under-sweep fixation duration be larger in scanning than in reading? There has been discussion that corrective saccades are driven by visual feedback following a saccade (Prablanc et al., 1978, Prablanc & Jeannerod, 1975) and it has also been reported that corrective saccade latencies in non-reading tasks are shorter when saccades land farther from their intended target (Becker, 1972). Given that participants' accurate line-initial fixations during scanning were closer to the left margin, it may be that scanning requires a more granular encoding strategy where participants target the very start of a line whereas they are targeting the preferred or optimal viewing location during reading (McConkie et al., 1989; Rayner, 1979), which is further from the left margin. If this were the case, retinal feedback during scanning would more rapidly indicate a deviation from the intended location of the saccade even when under-sweep landing positions were comparable; resulting in shorter under-sweep fixations.

To conclude, our research illustrates the remarkable consistency of the oculomotor system given that across both reading and scanning line-final and under-sweep fixations were shorter than intra-line fixations while accurate line-initial fixations were longer. While the basic pattern of return-sweep differences was observed for both tasks, there were nuanced differences. The reduction in line-final fixations did not differ across tasks, enabling us to conclude that shorter line-final fixation durations during reading cannot be attributed to higher-level lexical processing. Instead, it is likely driven by oculomotor or visual processing. By contrast, there was evidence that the increase in duration for accurate line-initial fixations was smaller during scanning, suggesting these fixations likely reflect a combination of both linguistic and oculomotor processes.

References

Abrams, S., & Zuber, B. (1972). Some temporal characteristics of information processing during reading. *Reading Research Quarterly*, 40–51.

Adedeji, V. I., Vasilev, M. R., Kirkby, J. A., & Slattery, T. J. (2022). Return-sweep saccades in oral reading. *Psychological Research*, *86*(6), 1804–1815.

Al-Zanoon, N., Dambacher, M., & Kuperman, V. (2017). Evidence for a global oculomotor program in reading. *Psychological Research*, *81*(4), 863–877.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Becker, W. (1972). The control of eye movements in the saccadic system. *Bibliotheca Ophthalmologica: Supplementa Ad Ophthalmologica*, *82*, 233–243.

Becker, W. (1976). Do correction saccades depend exclusively on retinal feedback? A note on the possible role of non-retinal feedback. *Vision Research*.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28.

Carr, J. W., Pescuma, V. N., Furlan, M., Ktori, M., & Crepaldi, D. (2022). Algorithms for the automated correction of vertical drift in eye-tracking data. *Behavior Research Methods*, *54*(1), 287–310.

Christofalos, A. L., Laks, M., Wolfer, S., Dias, E. C., Javitt, D. C., & Sheridan, H. (2024). Lower-level oculomotor deficits in schizophrenia during multi-line reading: Evidence from return-sweeps. *Quarterly Journal of Experimental Psychology*, 17470218231220752.

Dickey, J. M. (1971). The weighted likelihood ratio, linear hypotheses on normal location parameters. *The Annals of Mathematical Statistics*, 204–223.

Gagl, B., Gregorova, K., Golch, J., Hawelka, S., Sassenhagen, J., Tavano, A., Poeppel, D., & Fiebach, C. J. (2022). Eye movements during text reading align with the rate of speech production. *Nature Human Behaviour*, *6*(3), 429–442.

Hawley, T., Stern, J., & Chen, S. (1974). Computer analysis of eye movements during reading. *Literacy Research and Instruction*, *13*(4), 307–317.

Heller, D. (1982). Eye movement in reading. In R. Groner & P. Fraisse (Eds.), *Cognition and eye-movements* (pp. 139–154). Deutscher Verlag der Wissenschaften.

Hoaglin, D. C., & Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association*, *82*(400), 1147–1149.

Hofmeister, J. (1998). *Über korrektursakkaden beim lesen von texten und bei leseähnlichen aufgaben*. Shaker.

Hofmeister, J., Heller, D., & Radach, R. (1999). The return sweep in reading. *Current Oculomotor Research: Physiological and Psychological Aspects*, 349–357.

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye-movements in sentence and paragraph reading. *Quarterly Journal of Experimental Psychology*, *63*(9), 1838–1857.

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (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.

Nicenboim, B., Schad, D., & Vasishth, S. (2021). An introduction to bayesian data analysis for cognitive science. *Under Contract with Chapman and Hall/CRC Statistics in the Social and Behavioral Sciences Series*.

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

Parker, A. J., Kirkby, J. A., & Slattery, T. J. (2017). Predictability effects during reading in the absence of parafoveal preview. *Journal of Cognitive Psychology*, *29*(8), 902–911.

Parker, A. J., Kirkby, J. A., & Slattery, T. J. (2020). Undersweep fixations during reading in adults and children. *Journal of Experimental Child Psychology*, *192*, 104788.

Parker, A. J., Nikolova, M., Slattery, T. J., Liversedge, S. P., & Kirkby, J. A. (2019). Binocular coordination and return-sweep saccades among skilled adult readers. *Journal of Vision*, *19*(6), 10–10.

Parker, A. J., Räsänen, M., & Slattery, T. J. (2023). What is the optimal position of low-frequency words across line boundaries? An eye movement investigation. *Applied Cognitive Psychology*, *37*(1), 161–173.

Parker, A. J., & Slattery, T. J. (2019). Word frequency, predictability, and return-sweep saccades: Towards the modeling of eye movements during paragraph reading. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(12), 1614.

Parker, A. J., & Slattery, T. J. (2021). Spelling ability influences early letter encoding during reading: Evidence from return-sweep eye movements. *Quarterly Journal of Experimental Psychology*, *74*(1), 135–149.

Parker, A. J., & Slattery, T. J. (2024). Frequency and predictability effects for line-final words. *Journal of Experimental Psychology: Human Perception and Performance*.

Parker, A. J., Slattery, T. J., & Kirkby, J. A. (2019). Return-sweep saccades during reading in adults and children. *Vision Research*, *155*, 35–43.

Prablanc, C., & Jeannerod, M. (1975). Corrective saccades: Dependence on retinal reafferent signals. *Vision Research*, *15*(4), 465–469.

Prablanc, C., Massé, D., & Echallier, J. F. (1978). Error-correcting mechanisms in large saccades. *Vision Research*, *18*(5), 557–560.

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

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, *85*(3), 618.

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

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

Schroeder, S. (2019). popEye-an integrated r package to analyse eye movement data from reading experiments. *Journal of Eye Movement Research*, *12*(7).

Shebilske, W. L. (1976). Extraretinal information in corrective saccades and inflow vs outflow theories of visual direction constancy. *Vision Research*, *16*(6), 621–628.

Slattery, T. J., & Parker, A. J. (2019). Return sweeps in reading: Processing implications of undersweep-fixations. *Psychonomic Bulletin & Review*, *26*(6), 1948–1957.

Slattery, T. J., & Vasilev, M. R. (2019). An eye-movement exploration into return-sweep targeting during reading. *Attention, Perception, & Psychophysics*, *81*, 1197–1203.

Team, R. C. D. (2020). *R: A language and environment for statistical computing. 82.*

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for british english. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Vasilev, M. R., Adedeji, V. I., Laursen, C., Budka, M., & Slattery, T. J. (2021). Do readers use character information when programming return-sweep saccades? *Vision Research*, *183*, 30–40.

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

Voeten, C. C. (2023). *Buildmer: Stepwise elimination and term reordering for mixed-EffectsRegression*. https://CRAN.R-project.org/package=buildmer

Wagenmakers, E.-J., Verhagen, J., Ly, A., Matzke, D., Steingroever, H., Rouder, J. N., & Morey, R. D. (2017). The need for bayesian hypothesis testing in psychological science. *Psychological Science Under Scrutiny: Recent Challenges and Proposed Solutions*, 123–138.

Wang, M., Gill, D. E., Judge, J., Zang, C., Bai, X., & Liversedge, S. P. (2024). Column setting and text justification influence return-sweep eye movement behavior during chinese multi-line reading. *Cognitive Research: Principles and Implications*, *9*(1), 34.