

**Return-sweep saccades during reading in adults and children.**

Adam J. Parker<sup>1\*</sup>

Timothy J. Slattery<sup>2</sup>

Julie A. Kirkby<sup>3</sup>

Word Count: 6,773

Affiliations:

<sup>1</sup>: Department of Psychology, Faculty of Science & Technology, Bournemouth University, Talbot Campus, Poole, BH12 5BB. United Kingdom. Tel: 01202 965503, Email: [parkera@bournemouth.ac.uk](mailto:parkera@bournemouth.ac.uk).

<sup>2</sup>: Department of Psychology, Faculty of Science & Technology, Bournemouth University, Talbot Campus, Poole, BH12 5BB. United Kingdom. Tel: 01202 962556, Email: [tslattery@bournemouth.ac.uk](mailto:tslattery@bournemouth.ac.uk).

<sup>3</sup>: Department of Psychology, Faculty of Science & Technology, Bournemouth University, Talbot Campus, Poole, BH12 5BB. United Kingdom. Tel: 01202 965251, Email: [jkirkby@bournemouth.ac.uk](mailto:jkirkby@bournemouth.ac.uk).

\*Correspondence regarding this article should be addressed to Adam J. Parker, Department of Psychology, Faculty of Science & Technology, Bournemouth University, P104 Poole House, Talbot Campus, Poole, BH12 5BB. Tel: 01202 965503, Email: [parkera@bournemouth.ac.uk](mailto:parkera@bournemouth.ac.uk).

### **Abstract**

During reading, eye movement patterns differ between children and adults. Children make more fixations that are longer in duration and make shorter saccades. Return-sweeps are saccadic eye movements that move a reader's fixation to a new line of text. Return-sweeps move fixation further than intra-line saccades and often undershoot their target. This necessitates a corrective saccade to bring fixation closer to the start of the line. There have been few empirical investigations of return-sweep saccades in adults, and even fewer in children. In the present study, we examined return-sweeps of 47 adults and 48 children who read identical multiline texts. We found that children launch their return-sweeps closer to the end of the line and target a position closer to the left margin. Therefore, children fixate more extreme positions on the screen when reading for comprehension. Furthermore, children required a corrective saccade following a return-sweep more often than adults. Analysis of the duration of the fixation preceding the corrective saccade indicated that children are as efficient as adults at responding to retinal feedback following a saccade. Rather than consider differences in adult's and children's return-sweep behaviour an artefact of oculomotor control, we believe that these differences represent adult's ability to utilise parafoveal processing to encode text at extreme positions.

**Keywords:** Reading; Eye movements; Children; Return-sweep saccades; Oculomotor control.

**Highlights**

- Children launch their return-sweeps closer to the end of the line of text.
- Children prefer a viewing location closer to the start of the line.
- Children showed increased rates of return-sweep undershoot error (RUE).
- Children are just as efficient as adults in responding to RUE.

## 1. Introduction

Reading requires saccadic eye movements. Saccades are separated by brief pauses called fixations, during which readers encode information about the text. The eye movements of developing readers differ from those of adults. Children make more and longer fixations (Blythe, 2014; Blythe & Joseph, 2011; Reichle et al., 2013; Schroeder, Hyönä, & Liversedge, 2015). They also make shorter saccades (Blythe et al., 2006) and refixate words more frequently before leaving them (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Blythe, Häikiö, Bertram, Liversedge, & Hyönä, 2011; Joseph, Liversedge, Blythe, White, & Rayner, 2009). Additionally, children show a reduced perceptual span from which they extract information from the text (Rayner, 1986). These differences are observed despite children being as efficient as adults in the rate at which they extract visual information from the text (Blythe et al., 2009, 2011).

We know much more about children's eye movements occurring within a line of text than we do about eye movements that cross line boundaries. Return-sweeps are saccadic eye movements that move a reader's fixation from the end of one line to the start of the next. During natural reading, return-sweeps precede or follow approximately 20% of all fixations (Rayner, 1998). The prevalence of these saccades is highlighted when considering an example. For instance, the first page of *Harry Potter and the Philosopher's Stone* requires readers to make 30 return-sweeps without any line re-reading. This translates as approximately 6,500 return-sweeps for the 233 pages in the 1997 Bloomsbury paperback edition. Without a direct examination of return-sweeps during children's processing of multiline texts it is difficult to fully quantify differences in eye movements between adult and child readers as they develop from novice to expert.

### 1.1 Developmental aspects of eye movements during reading.

While several saccadic parameters (i.e. velocity, adaptation) develop throughout early childhood (Bucci & Seassau, 2012; Doré-Mazars, Vergilino-Perez, Lemoine, & Bucci, 2011; Salman et al., 2006), others (i.e. precision) show a developmental curve and stabilised during adulthood (Luna, Velanova, & Geier, 2008). With regards to where adult and child readers move their eyes, these two populations appear similar. With little reading experience, children aged 6- to 7-years-old target saccades towards word centres (McConkie et al., 1991). The initial landing position within a word is similar between adults and children aged 7- to 11-years-old, with landing positions varying with word length. However, 7- to 11-year-olds have been found to make more refixations than adults (Joseph et al., 2009). Despite some differences, the rate of visual processing does not differ between adults and children. Evidence from the disappearing text paradigm (Blythe et al., 2009, 2011; Rayner, Liversedge, White, & Vergilino-Perez, 2003) indicates that 7-year-old children can extract the visual information that is necessary for linguistic processing within 60 ms. Together these findings indicate that adults' and children's eye movements are similar with respect to where they target words and the rate at which visual encoding occurs.

Like adults, children extract information that is right of fixation. The window from which 7- to 9-year-olds extract information is estimated to extend from 3- to 4-character spaces to the left of fixation to 11-letters to the right; while the perceptual span of 11-year-old children is estimated to extend from 3- to 4-character spaces to the left of fixation to 14-characters to the right (Häikiö, Bertram, Hyönä, & Niemi, 2009; Rayner, 1986, Sperlich, Schad, & Laubrock, 2015). This asymmetry is comparable to that in skilled adult readers who perceive information from a window that extends 3- to 4-characters from the right of fixation to 14- to 15-characters to the right (see Rayner, 1998, 2009 for reviews). A similar pattern has been observed for the letter identification span in children (Häikiö et al., 2009). Thus, it

would appear that with very little reading instruction, children deploy attentional resources to the right of fixation.

The spatial extent of the effective field of vision only becomes adult-like at 11-years old (Häikiö et al., 2009; Rayner, 1986). Because the rate of visual encoding is similar between adults and children (Blythe et al., 2009, 2011), differences in fixation time measures are thought to reflect increased demands of foveal processing in children. As a result, it is argued that children allocate less attention to the upcoming word. In skilled adult readers, parafoveal processing decreased with increasing text difficulty (Rayner, 1986; Henderson & Ferreira, 1990; White, Rayner, & Liversedge, 2006), and reading skill (Chace, Rayner, & Well, 2005). Thus, the developmental changes in parafoveal likely reflect increased linguistic processing efficiency. Studies directly examining children's parafoveal processing indicate that, despite having a reduced perceptual span relative to adults, children process information to the right of fixation (Pagán, Blythe, & Liversedge, 2016; Häikiö, Bertram & Hyönä, 2010; Marx, Hawelka, Schuster, & Hutzler, 2015, 2017; Marx, Hutzler, Schuster, & Hawelka, 2016). Tiffin-Richards and Schroeder (2015) reported evidence to suggest a developmental shift in the type information extract from parafoveal vision. While children showed a phonological preview benefit, adults did not. Instead, adults showed a transposed letter effect indicating that adults use orthographic information in parafoveal vision. However, studies involving readers of English have reported that, like adults, children extract letter identity and position in parafoveal vision (Pagán et al., 2016). It is possible that orthographic depth may modulate differences between scripts.

Similarities have also been observed in adult's and children's foveal processing. Longer words are the recipients of more and longer fixations (Blythe et al., 2010; Hyönä & Olson, 1995; Joseph et al., 2009), with these length effects being more pronounced in children (Joseph et al., 2009). Word length effects appear to decrease in magnitude between

7- and 9-years of age (Huestegge, Radach, Corbic, & Huestegge, 2009), and decrease with increasing reading proficiency (Hyönä & Olson, 1995). Like adults, fixation durations of children tend to be longer for infrequent words (Blythe et al., 2009; Hyönä & Olson, 1995). Again, these effects are more pronounced for children, especially when age-appropriate frequency norms are used (Joseph, Nation, & Liversedge, 2013).

The reviewed studies indicate both similarities and differences in the eye movements of adult and child readers. Simulations using computational models of eye movement behaviour (e.g. Engbert & Kliegl, 2011; Reichle, 2011) have attempted to explain such differences. By adjusting parameters which account for saccadic programming and execution and/or linguistic processing efficiency, Reichle et al. (2013) were able to examine several predictions regarding how changes in oculomotor, visual, and linguistic processes shape the developmental trajectory. While adjustments to parameters associated with saccadic programming and execution were not sufficient to simulate eye movement differences between adults and children, changes in the rate of linguistic processing were.

In sum, adults and children are remarkably similar in terms of their oculomotor control. Instead, it appears that these developmental changes result from changes in the rate of lexical processing. A unifying explanation is developing readers must allocate more of their attentional and linguistic processing resources to the fixated word when reading for comprehension. Consequently, developing readers are less capable to engage in the processing of extrafoveal information. With regards to return-sweep saccades such an account would predict differences in launch sites and landing positions between adults and children. Specifically, compared to adults, children would launch return-sweeps closer to the ends of lines and need to fixate closer to the start of a new line before beginning their left-to-right reading pass of the line.

### **1.2 Return-sweep saccades during skilled reading.**

Return-sweep saccades are to fixate the next line of text. They are typically launched close to the end of the line and take 30-125 ms to complete. The fixation following a return-sweep is generally 5- to 7-characters from the leftmost letter of a line. For intra-line saccades, readers tend to land at preferred viewing location (PVL), which is usually left of word centres (Rayner, 1979). Word centres are represent the optimal viewing position (OVP) as this is where word recognition processes are most efficient (O'Regan & Jacobs, 1992; O'Regan, Lévy-Schoen, Pynte, & Brugailière 1984). Readers target a position close to the OVP because they are able to use spaces between words in parafoveal vision for saccade targeting (e.g. Morris, Rayner, & Pollatsek, 1990; Pollatsek, Juhasz, Reichle, Machacek, & Rayner, 2008). Unlike intra-line saccades, the target of a return-sweep lies far beyond parafoveal vision. Therefore, it is unlikely that readers are able to target return-sweeps towards the centre of the line-initial word given visual acuity limitations in the periphery. Indeed, when Heller and Radach (1992) monitored participants eye movements as they read a classic novel in which they estimated the OVP to be 3- to 4-letters from the left margin, return-sweeps had their peak significantly further to the right. Heller and Radach (1993; as cited in Hofmeister et al., 1999) subsequently reported that return-sweep landing position was independent of the length of the line-initial word. Together, these results indicate that return-sweep targeting does not rely on properties of line-initial words.

Like all saccades, return-sweeps are the result of muscular contractions and are subject to error. Because return-sweeps move the eyes a greater distance than other reading saccades, they will be influenced more by this error (McConkie, Kerr, Reddix, & Zola, 1988). This means that return-sweeps will often fall short of their target—the start of a new line— requiring a corrective saccade (Heller, 1982; Hofmeister et al., 1999; Parker, Kirkby, & Slattery, 2017; Rayner, 1998). The likelihood of initiating a corrective saccade increases the further readers land from the left margin (Radach & Heller, 1993). Applied questions

motivate early work. For example, Tinker and colleagues examined the effects of typographic manipulations on return-sweep accuracy (Tinker, 1963). While these endeavours showed that return-sweep accuracy was related to line length (see Hofmeister et al., 1999 for a discussion), they did not detail how return-sweeps differ between skilled and developing readers. With regards to how adults and children differ in relation to their return-sweep behaviour, research is limited. Netchine, Guihou, Greenbaum, and Englander (1983) found that, compared to adults, children's return-sweeps were more likely to fall short of their target and require a corrective saccade toward the left margin when reading. Dyslexic readers exhibit a similar pattern when reading aloud (Trauzettle-Klosinski, et al., 2010). Together these results show that reading ability influences, at the very least, the accuracy of return-sweep saccades.

### **1.3 Reading fixations and return-sweeps**

For the purpose of examining return-sweeps in reading, it is useful to define specific populations of reading fixations. As previously noted, ~20% of all reading fixations are preceded or followed by a return-sweep. Therefore, ~80% of fixations belong to a "intra-line" reading fixation population which are non-adjacent to return-sweeps (see Rayner, 1998; 2009 for reviews). The fixations that are adjacent to return-sweeps fall into one of three distinct groups (line-final, accurate line-initial, and undersweep). Line-final fixations are those that immediately precede a return-sweep. Accurate line-initial fixations immediately follow a return-sweep and are themselves followed by a rightward reading saccade. That is, this population of accurate line-initial fixations landed close enough to target of the return-sweep that the reader was able to begin their normal left to right (for readers of English) reading pass of that line of text. Undersweep-fixations are those that immediately follow a return-sweep and are themselves followed by a leftward corrective saccade. That is, this population

of undersweep-fixations fell short of the return-sweep's target requiring the reader to make an additional leftward saccade prior to beginning their left to right reading pass of the line.

Line-final fixations are typically shorter than intra-line reading fixations (Rayner, 1977; see also Abrams & Zuber, 1972; Hawley, Stern, & Chen, 1974). Furthermore, fixation durations decrease as readers progress towards the end of lines in multiline texts (Kuperman, Dambacher, Nuthmann, & Kliegl, 2010). Rayner (1977) suggested that this decrease in fixation duration presents a lack of parafoveal processing for the upcoming word. Alternatively, Kuperman et al. (2010) and Mitchell, Shen, Green, and Hodgson (2008) attributed this reduction to oculomotor programming. Consistent with this oculomotor account Abrams and Zuber (1972) argued that shorter line-final fixations result from return-sweep planning. Furthermore, Hofmeister (1997) reported that following a 50% degradation of the text there was a 20 ms increase in duration for all reading fixations other than line-final fixations, suggesting that line-final fixations are relatively uninvolved in linguistic processing. This, at the very least, suggests that skilled adult reader's fixation durations are less effected by stimuli quality when they immediately precede a return-sweep. To date, only one study that has investigated children's eye movement behaviour as the end of the line. Tiffin-Richards and Schroder (2018) reported that with increasing age, children displayed shorter reading times on line-final words. At grade 2, readers did not show a reduction in gaze duration for line-final words. By grade 4, developing readers showed a reduction that did not differ from adults. A similar effect was noted for first-fixation duration. While this shows an overall reduction in reading time at the end of the line, it is important to consider what these measures represent. Neither one explicitly examines line-final fixations which for children will often be the second or third fixation on a word. Therefore, it is difficult to conclude the whether the last fixation, specifically, is shorter in developing readers relative to adults.

Accurate line-initial fixations, those that reach their intended location without requiring a corrective saccade, tend to be longer than intra-line reading fixations by around 30-50 ms (Rayner, 1977). Several explanations for longer fixation times following an accurate return-sweep exist. Stern (1981) argued that this is likely to result from binocular coordination processes. That is vergence movements and a period of reorientation are required to resolve increased binocular disparity following return-sweep execution prior to the processing of new information on a line. However, this vergence explanation is, difficult to discern from an account which attributes the increased fixation duration following an accurate return-sweeps to a lack of parafoveal preview for the new line. Such an explanation posits that because readers cannot parafoveally process the information at the start of a new line on the fixation prior to the return-sweep, the fixation times on these words are substantially longer. Consistent with this, Parker et al. (2017) noted that gaze durations were shorter for words in the middle of a line where they could be pre-processed during the prior fixation, than when they were the first word on a new line, where preview during the prior fixation is denied.

A significant proportion of return-sweeps are followed by an immediate leftward corrective saccade that brings the eyes closer towards the left margin (e.g. Hofmeister et al., 1999). In such cases, the intervening fixation tends to be ~120 ms shorter than intra-line reading fixations (130 ms; Heller, 1982). Parker et al. (2017) have termed these “undersweep-fixations”. It has long been assumed that these very brief undersweep-fixations are not involved in on-going linguistic processing (Hawley et al., 1974; Shebilske, 1975). Instead they have been considered an oculomotor response based on retinal feedback which represents the deviation of a return-sweep landing position from its intended goal (Becker, 1976; Hofmeister et al., 1999). Consistent with this, Heller and Radach (1993) have reported that undersweep-fixations are terminated as a function of the distance that a return-sweep

lands from the left margin rather than by aspects of the text itself (i.e. line length) or the distance traversed by the return-sweep. What currently remains unclear is whether the saccade following an undersweep-fixations is predominantly inter-word or intra-word. Children's increased prevalence to initiate a corrective saccade following a return-sweep means that they will have more undersweep-fixations and fewer accurate line-initial fixations than adult readers. Understanding the role these fixations play during reading is therefore one of the goals of the present research.

#### **1.4 The present study.**

To provide a detailed account of return-sweep saccade differences between developing and skilled adult readers, we examine the return-sweep behaviour of 47 adults and 48 children. Both adults and children read the same 20 passages of text which were specifically constructed to allow for their comparison across these groups of readers. Below we set out our specific predictions in relation to return-sweep and corrective saccade parameters and fixation durations.

In comparison to adults, developing readers allocate more processing resources to the fixated word. As a result, children appear less efficient at processing extrafoveal information. Indeed, Häikiö et al. (2009) reported data showing that the letter identification span is less efficient in developing readers. Therefore, we predict that children will fixate closer to the end of a line of text to encode more information foveally. Similarly, we expect that children will prefer a line initial viewing location closer to the left margin. Based on prior work (e.g. Netchine et al., 1983), we predict that children will make more undersweep-fixations than adults.

For fixation durations, we make specific predictions for each fixation population (intra-line, line-final, accurate line-initial, and undersweep). As has been shown repeatedly, we expect that intra-line reading fixations will be shorter for adults than for children. In

adults, it is well documented that line-final fixations are shorter than standard reading fixations, we expect to replicate this result with the adult data. Assuming that line-final fixations are shorter as a consequence of return-sweep planning as opposed to lexical processing, we predict that children will show a similar reduction in line-final fixation durations as adults. If, however, the reduction in duration for line-final fixations reflects a lack of parafoveal processing, then it would be expected that children's line-final fixations would not differ from their intra-line fixations due to their reliance on foveal processing. For accurate line-initial fixations, we expect these to be longer than intra-line reading fixations. Prior to a return-sweep, the words at the start of the next line are unavailable for parafoveal pre-processing and their identification requires foveal processing. Therefore, we expect adult's accurate line-initial fixations to be longer than their intra-line reading fixations. However, since children's intra-line reading fixations are largely involved with foveal processing we expect that these would not differ from their accurate line-initial fixations. Finally, with regards to undersweep-fixations we expect these to be shorter than any other fixation population for both adults and children due to their being under oculomotor control rather than lexical control.

## 2. Method

### 2.1 Participants.

Fifty-two children, aged 6- to 9-years old<sup>1</sup>, were contacted through local schools. Children were contacted with the consent of parents and headteachers who were informed on the study's purpose. Children were native English speakers, naïve to the study's purpose, and had normal or corrective-to-normal vision. One child did not complete the eye movement study and three were excluded: two performed below chance on the reading comprehension questions (< 50% accuracy) and one performed poorly on the IQ measures. This left 48 children in the final sample (23 male; mean age: 7.5 years,  $SD = .85$  years).

Fifty-three adult participants, aged 18- to 32-years old, were all members of the Bournemouth University community and provided informed written consent. Adults were native English speakers, naïve to the study's purpose, and had normal or corrective-to-normal vision. Two adults did not complete the eye movement study and a further four adults were excluded; two performed below chance on the reading comprehension questions and two performed poorly on the IQ measures. This left 47 adults in the final sample (15 male; mean age: 21.0 years,  $SD= 2.60$  years).

## **2.2 Off-line ability measures.**

Prior to the eye movement study, participants completed off-line measures assessing reading and general ability. Reading efficiency was assessed using the Test of Word Reading Efficiency 2 (TOWRE; Torgesen, Wagner, & Rashotte, 2011). Within 45 seconds, participants read aloud as many words and non-words from a list as possible. The letters and digits subscale of the Rapid Automated Naming and Rapid Alternating Stimulus Test (RAN/RAS; Wolf & Denckla, 2005) was used to assess participants' ability to perceive and name a visual symbol. IQ was assessed using two subtests of the Wechsler Abbreviated Scale of Intelligence II (WASI-II; Wechsler, 2011): a) the vocabulary subtest; b) the matrix reasoning subtest. Mean performance was at an age-appropriate level on all measures (see Table 1). Importantly, children's mean reading age was 9.5 years ( $SD= 2.28$ ; range: 7.00-16.25) indicating that children were reading at a level that was sufficient for the current study.

Table 1. Mean performance of participants included in the experiment on off-line tasks.

	Adults	Children	Test mean
TOWRE	109.6 (11.03)	109.3 (13.07)	100
RAN/RAS	110.4 (4.59)	108.3 (11.68)	100
WASI-II Vocabulary	53.9 (8.91)	53.8 (10.80)	50
WASI-II Matrices	50.7 (8.42)	48.1 (9.12)	50

*Note.* Mean standardised scores are shown for the TOWRE and RAN/RAS. Mean *T* scores are shown for the WASI-II subscales. Standard deviations are shown in parentheses. There were no significant differences between adults and children on any of the off-line tasks (all *ps* > .098).

### 2.3 Materials.

Adults and children read the same 20 passages of text to avoid differences in materials confounding group differences. Passages were taken from *Harry Potter and the Philosopher's Stone* and were edited to ensure that they were suitable for the youngest readers. To accomplish this, we first acquired Age of Acquisition (AoA) ratings for each word in a passage (from Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012). We then replaced words with an AoA above 7 so long as this did not result a change of meaning for the passage. For example, item 2 in the supplemental materials originally read: *The repaired alarm clock rang at six o'clock the next morning.* The word “repaired” has an AoA rating of 11.95 years. This was substituted with the word “enormous” which had an AoA rating of 6.53 years. Across all 20 items, there were 7 which required this replacement procedure<sup>2</sup>.

Text was displayed across two-to-three lines (see Figure 1). Each passage contained 10 to 17 words ( $M= 13.35$  words). Words varied from 1 to 11 letters ( $M= 4.96$  letters) and had an average Zipf frequency (van Heuven, Mandera, Keuleers, & Brysbaert, 2014) based on the SUBTLEX database (Brysbaert & New, 2009) of 5.36 (range: 1.81-7.67). For

comparison, the average Zipf frequency for children, based on the CBBC subtitle frequency, was 5.37 (range: 2.17-7.57).

Harry walked round and round his new room.  
Someone knew he had moved out of his  
cupboard.

Figure 1. Example stimuli read by both adults and children. All stimuli required readers to make at least one return-sweep saccade.

#### **2.4 Apparatus.**

Movements of the right eye were recorded at a rate of 1000 Hz using an SR Research EyeLink 1000 plus eye-tracker. Text was presented on a BenQ XL2410 T LCD monitor with a 1920 x 1080 resolution at a distance of 80 cm. Each character was presented in 22-point Courier New font so that each character equated to approximately  $0.3^\circ$  of horizontal visual angle.

#### **2.5 Procedure.**

The procedure was approved by Bournemouth University's Research Ethics Code of Practice and in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Children were tested within schools in rooms suitable for eye tracking. Adults were tested in laboratory rooms at Bournemouth University. Participants first completed the off-line ability measures. Next, participants completed the eye movement study in which they sat in front of a computer monitor. Chin and forehead rests were used to ensure comfort and minimise head movements. Participants undertook a 9-point calibration

and validation procedure. If the average error was greater than .3 degrees of visual angle or greater than 1 degree at any individual point, this procedure was repeated. Text appeared contingent on a stable fixation being detected in the top left of the screen. Participants read silently for meaning and answered occasional comprehension questions. They were instructed to press a button on the keyboard when they had finished reading each text. Participants responded to TRUE/FALSE comprehension questions following seven passages (35%) by pressing one of two buttons on the keyboard. For the example stimuli shown in Figure 1, the accompanying comprehension question was:

*Harry walked around his new room.*

*TRUE/FALSE*

Items were presented in random order. Each participant read a total of 20 passages. Passages were embedded in a larger list of stimuli containing four practice passages and 40 fillers. The eye-tracking session lasted approximately 20-30 minutes. Testing lasted approximately 1 hour.

## **2.6 Data Analysis.**

For each passage of text, we considered each fixation. Fixations less than 80 ms within a character space of a previous or subsequent fixation were combined with that fixation. Other fixations shorter than 80 ms and greater than 1200 ms were excluded prior to analysis. Trials in which there was tracker loss or 5 or more blinks were also removed. Trials were excluded if a blink preceded or followed a return-sweep. Analysis was conducted on the remaining 93.4% of data (30,797 fixations).

Linear mixed effects (LME) models were fit to eye movement data within the R computing environment (version 3.5.1.; R Development Core Team, 2018) using the *lmer* function from the *lme4* package (version 1.1-18; Bates et al., 2018). As an estimate of effect size relative to the intercept, we report regression coefficients (*b*). Standard errors (*SE*) and *t*-

values are also reported. The lmerTest package (version 3.0-1; Kuznetsova, Brockhoff, & Christensen, 2017) was used to compute  $p$ -values. We consider cases where  $|t| > 1.96$  as statistically significant (Baayen, Davidson, & Bates, 2008). Generalized linear mixed effects (GLME) models were fit to binary dependent variables, such as undersweep-fixation likelihood, using the *glmer* function from lme4. The Wald  $z$  and its associate  $p$ -value is reported.

Models initially adopted the full random structure for both participants and items, with random intercepts and slopes (Barr, Levy, Scheepers, & Tily, 2013). If models did not converge, the random structure was simplified to reduce overfitting (Bates, Kliegl, Vasishth, & Baayen, 2015). Models were simplified by removing random effects that were perfectly or near-perfectly correlated with others so long as model fit was not reduced.

### 3. Results

All participants scored at least 50% on the comprehension questions (adults:  $M=87.0\%$ ,  $SD=9.81$ ,  $range=71-100\%$ ; children:  $M=81.6\%$ ,  $SD=13.47$ ,  $range=57-100\%$ ). Both adults,  $t(47)=25.50$ ,  $p<.001$ , and children,  $t(47)=16.25$ ,  $p<.001$ , scored significantly higher than chance with adults making significantly more accurate responses,  $t(86)=2.11$ ,  $p=.037$ ,  $d=.46$ . Below we report return-sweep and corrective saccade parameters and fixation durations for each fixation population. Because skilled adult reading represents the end point of reading development, we used adult data as the baseline (intercept) for all models with a single predictor. While regression coefficients are reported in the main body of text, (G)LME results tables are included in the supplemental materials.

#### 3.1 Return-sweep and corrective saccade parameters.

We examined three saccade parameters: return-sweep launch site (the number of characters from the end of a line that a return-sweep was launched), return-sweep landing position (the number of characters from the left margin of the new line that the return-sweep

landed), and the frequency of undersweep-fixations (the percentage of return-sweeps that required a corrective saccade). These metrics were considered for 2,083 return-sweeps (1,079 adult return-sweeps) and 1,772 line initial fixations (953 adult line initial fixations). Mean return-sweep and corrective saccade parameters are shown in Table 2. For all analyses, an undersweep-fixation refers to a pause following a return-sweep that is immediately followed by a leftwards saccade towards the left margin. These cases reflect a mixture of within-word refixations and intra-word regressions. The majority of undersweep-fixations were followed by an inter-word regression (adults: 64%; children: 58%).

Table 2. Descriptive statistics for return-sweep and corrective saccade parameters between adults and children.

	Return-sweep launch site	Return-sweep landing position	Frequency of undersweep-fixations (%)
Adults	4.5 (2.58)	5.1 (2.46)	51.5 (50.01)
Children	3.3 (2.37)	4.7 (3.09)	62.37 (48.48)

*Note.* Return-sweep launch sites, amplitude and landing site are shown in character spaces.

Means are displayed with standard deviations in parenthesis.

First, we fit an LME model to return-sweep launch site data. The distribution of return-sweep launch sites is shown in Figure 2. Note that children very frequently launched their return-sweeps from a position close to the end of the line. The percentage of return-sweeps launched from the line final word are shown in Table 3. Prior to analysis, to exclude the extended right tail of the distribution, we excluded launch sites over 10 characters from the end of the line (8.6% of fixations). The model included a categorical fixed effect which coded whether participants were adults or children and included the full random structure ( $lmer(dv \sim group + (1 | participant) + (1 + group | item))$ ). This analysis indicated that

children's return-sweeps were executed closer to the end of the line than adults,  $b = -1.20$ ,  $SE = .19$ ,  $t = -6.19$ ,  $p < .001$ .

Table 3. Percentage of return-sweep launched from line final words, the percentage of which followed a refixation on a line final word, and whether the word had been fixated.

	Launched from line final word	Launched after a line final word	Line final word fixated	Line final word not fixated
Adults	67.1	12.1	75.0	25.0
Children	86.6	31.2	92.1	7.9

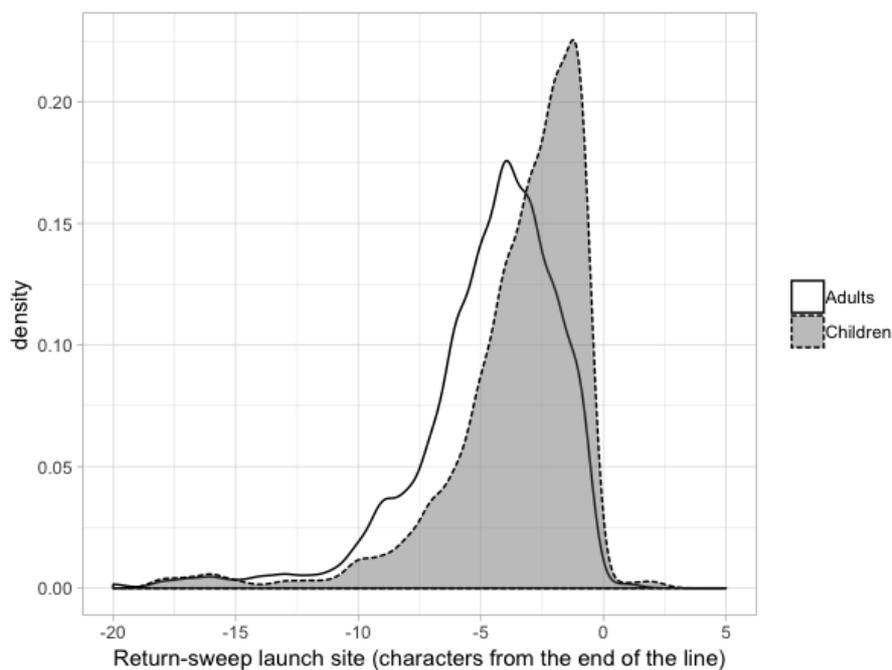


Figure 2. The distribution of return-sweep launch site for adults and children. On this scale zero refers to the end of the line. Minus values on the x-axis indicate that return-sweeps were made prior to the end of the line. Positive values on the x-axis indicate return-sweeps were made passed the end of the line.

To determine how line initial landing position differed between adults and children, we fit an LME model to landing site data. To account for the two distinct types of fixations that follow return-sweeps, we included an additional fixed effect that coded whether the line initial fixation was accurate or an undersweep-fixation (the distribution of landing position is shown in Figure 3 for adults and children). Prior to analysis, we excluded return-sweeps which landed more than 15 characters from the start of the line (1.2% of fixations). The final model included fixed effects for participant group, fixation type (accurate or undersweep), and their interaction. The model included the full random structure ( $lmer(dv \sim undersweep * group + (1 + undersweep | participant) + (1 + undersweep * group | item))$ ). Children's accurate return-sweeps landed closer to the left margin than did adults,  $b = -.98$ ,  $SE = .32$ ,  $t = -3.01$ ,  $p = .004$ . Adults' undersweep-fixations landed farther from the left margin than did accurate line-initial fixations,  $b = 2.26$ ,  $SE = .20$ ,  $t = -11.39$ ,  $p < .001$ . The interaction between group and undersweep likelihood indicated that the shift between landing sites of accurate line-initial and undersweep-fixations was not different for adults and children,  $b = .37$ ,  $SE = .28$ ,  $t = 1.29$ ,  $p = .203$ .

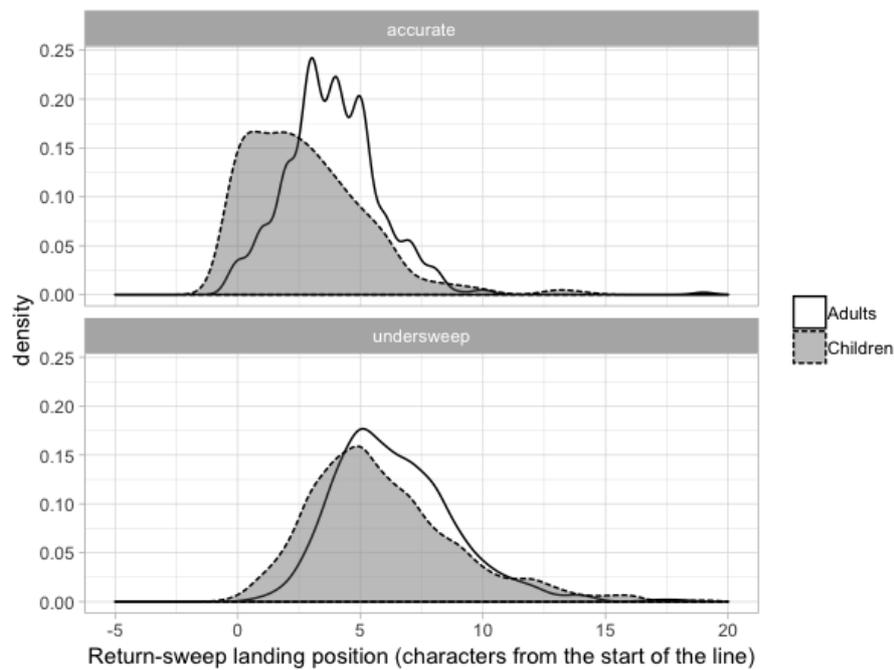


Figure 3. The distribution of return-sweep landing position for adults and children. On this scale zero refers to the start of the line. Minus values on the x-axis indicate that return-sweeps landed before the start of the line. Positive values on the x-axis indicate return-sweeps landed to the right of the left margin.

Finally, we fit a GLME model to predict undersweep-fixation likelihood. The model included fixed effects for participant group, landing site, and their interaction. The model included random intercepts for participants and items with random slopes for landing position across participants ( $glmer(dv \sim landingpos * group + (1 + landingpos | participant) + (1 | item))$ ). Landing site was included as prior research has shown that the likelihood of making a corrective saccade increases as return-sweeps land farther from the left margin. Following our previous analysis, we excluded data for return-sweeps which landed more than 15 characters away as these were very rare (0.9% of fixations). Children made more undersweep-fixations than adults,  $b = 3.68$ ,  $SE = .34$ ,  $z = 3.80$ ,  $p < .001$ . Line initial fixations were increasingly followed by a corrective saccade the further away that they landed from the left margin,  $b = 2.70$ ,  $SE = .09$ ,  $z = 11.22$ ,  $p < .001$ . The interaction between group and landing position (see Figure 4) indicated that, relative to adults, children were more likely to initiate a

corrective saccade from a location closer to the left margin,  $b = -.78$ ,  $SE = .11$ ,  $z = -2.35$ ,  $p = .019$ .

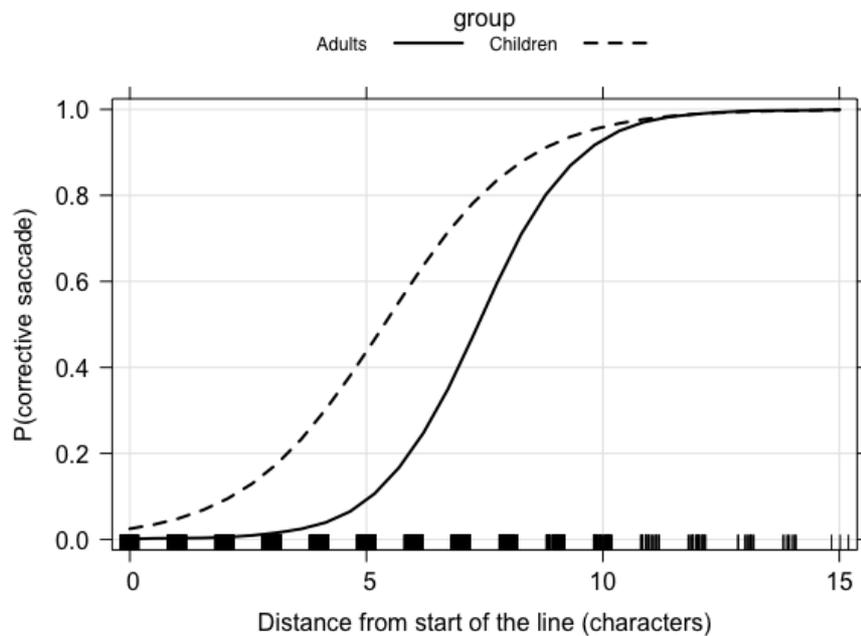


Figure 4. Probability of making a corrective saccade following a return-sweep as a function of distance from the start of the line and participant group. Predictions derived from the GLME are shown for adults (solid line) and children (dashed line).

### 3.2 Return-sweeps fixation durations.

We compared return-sweep fixation durations to intra-line reading fixations by coding a categorical variable with the following contrast values. Intra-line reading (non-return-sweep) fixations were coded as -1. Line-final fixations were coded as -0.5, accurate line-initial fixations as 0.5, and undersweep-fixations as 1. This coding scheme meant that intra-line reading fixations would represent the intercept to which the other “return-sweep fixations” were compared. To assess how fixation types differed between adults and children, an additional fixed effect for participant group was included and allowed to interact with fixation type. Means for each fixation population are shown in Figure 5.

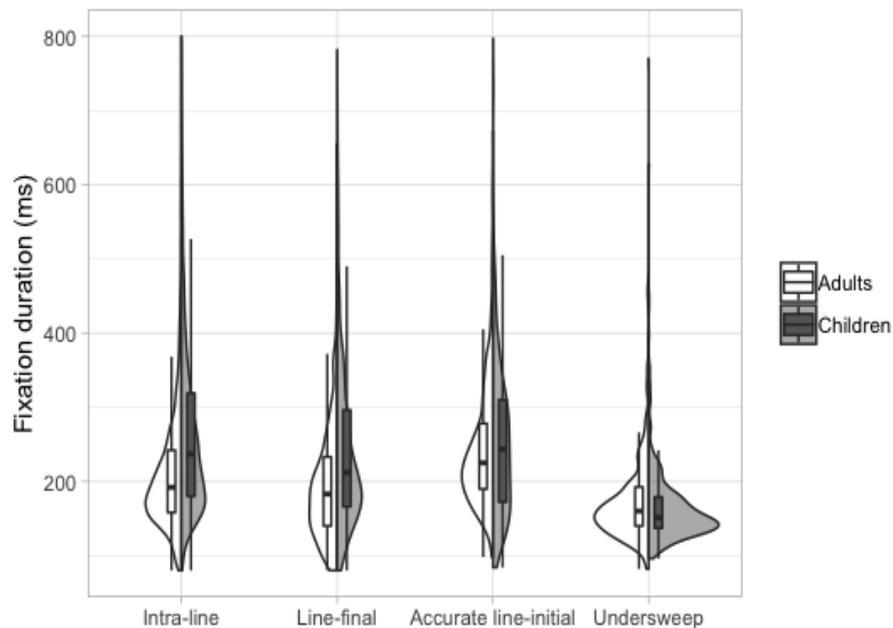


Figure 5. Split violin plot for fixation duration (ms) as a function of fixation population and participants group. The distribution of fixation duration for each fixation population is shown in white for adults and grey for children. Boxplots show the first quartile, median, and third quartile per fixation population. Though the y-axis set to a maximum of 800 ms, trimming procedures for analysis used 1,200 ms as an upper bound.

The model fit to log-transformed fixation duration data included random slopes for fixation population for participants and items and group for items ( $lmer(dv \sim \text{fixtype} * \text{group} + (1 + \text{fixtype} | \text{participant}) + (1 + \text{fixtype} + \text{group} | \text{item}))$ ). For adults, line-final fixations,  $b = -.03$ ,  $SE = .01$ ,  $t = -2.73$ ,  $p = .009$ , and undersweep-fixations,  $b = -.07$ ,  $SE = .01$ ,  $t = -6.10$ ,  $p < .001$ , were shorter than intra-line reading fixations, while accurate line-initial fixations were longer,  $b = .08$ ,  $SE = .01$ ,  $t = 6.38$ ,  $p < .001$ . On average, children's intra-line fixation durations were longer than adults,  $b = .10$ ,  $SE = .01$ ,  $t = 8.79$ ,  $p < .001$ . The difference in fixation duration for intra-line fixations and line-final fixations were similar between adults and children,  $b = -.01$ ,  $SE = .01$ ,  $t = -.53$ ,  $p < .601$ . The difference between intra-line fixations and accurate line-

initial fixations,  $b = -.09$ ,  $SE = .02$ ,  $t = -5.19$ ,  $p < .001$ , and undersweep-fixations,  $b = -.12$ ,  $SE = .01$ ,  $t = -8.23$ ,  $p < .001$ , differed between adults and children. While adults showed longer accurate line-initial fixations relative to intra-line fixations, children were numerically shorter. The reduction in fixation for undersweep-fixations relative to intra-line fixations was greater for children than adults<sup>3</sup>.

To directly assess how this pattern of fixations differed for children, we fit a model to log-transformed fixation duration data for children. The model included a fixed effect for fixation population and the full random structure: ( $lmer(dv \sim fixtype + (1 + fixtype / participant) + (1 + fixtype / item)$ ). Relative to intra-line fixations, line-final fixations were shorter,  $b = -.04$ ,  $SE = .01$ ,  $t = -2.71$ ,  $p = .011$ , as were undersweep-fixations,  $b = -.19$ ,  $SE = .01$ ,  $t = -15.13$ ,  $p < .001$ . Accurate line-initial fixations did not statistically differ from intra-line fixations,  $b = -.02$ ,  $SE = .02$ ,  $t = -.99$ ,  $p = .331$ .

The finding that children's accurate line-initial fixations did not differ in duration from intra-line fixations was unexpected. One candidate explanation is that adults use parafoveal information at the very start of the line to plan their subsequent saccades over the line (Kuperman et al., 2010). Children, however, may not engage in this saccade planning as they largely rely on foveal processing. Assuming that saccade planning would require a longer line-initial fixation, such logic would predict similar durations between intra-line and accurate line-initial fixations in children. To assess such a possibility, we fit an additional LME model to saccade amplitude. Here, saccade amplitude was used as an index of subsequent parafoveal processing. The model ( $lmer(dv \sim fixtype * Group + (1 + fixtype / participant) + (1 + fixtype / item)$ ) indicated that children made shorter intra-line forward (rightward) saccades than adults,  $b = -1.47$ ,  $SE = .26$ ,  $t = -5.68$ ,  $p < .001$ . Adults' saccade amplitudes were longer following an accurate line-initial fixation,  $b = .71$ ,  $SE = .30$ ,  $t = 2.38$ ,  $p = .021$ . The difference between saccade lengths following intra-line and accurate line-initial

fixations was smaller for children in comparison to adults,  $b = -1.10$ ,  $SE = .36$ ,  $t = -3.11$ ,  $p = .003$ .

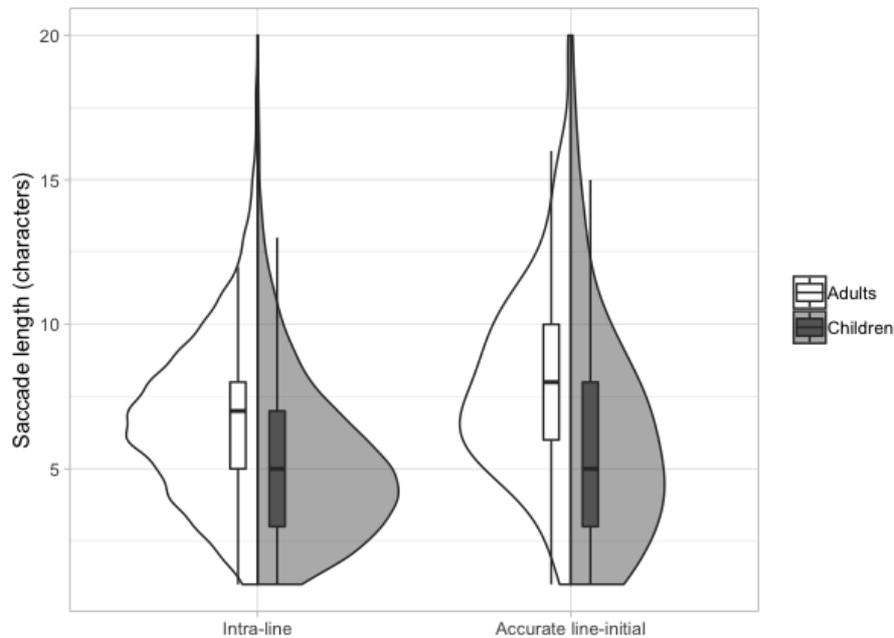


Figure 6. Split violin plot for following saccade length (characters) as a function of fixation population (intra-line and accurate line-initial) and participant group. The distribution of saccade length for each fixation population is shown in white for adults and grey for children. Boxplots show the first quartile, median, and third quartile per fixation population.

#### 4. Discussion

To examine return-sweep and corrective saccade parameters between adults and children, we compared their eye movements as they silently read multiline texts. We found, in accordance with previous research (Hofmesiter et al., 1999), that adult's launched return-sweeps relatively close to the end of the line, landed relatively close to the start of the next line, and frequently required corrective saccades to bring their fixation point closer to the start of the line following a return-sweep. Adults also showed a pattern of fixation durations reported in the literature: accurate line-initial fixations were longer than intra-line reading

fixations while line-final and undersweep-fixations were shorter. Together these results indicate that intra-line fixations and those adjacent to return-sweeps are differentially affected by linguistic processing.

The main focus of the present study was to compare children's oculomotor behaviour over multiline texts with that of adults. While intra-line saccade targeting has been shown to be remarkably similar between adults and children (cf. Joseph et al., 2009), the similarities and differences in return-sweep parameters and fixation durations had not been previously examined in a single study. Consistent with our predictions, we found that children preferred viewing locations at more extreme locations on lines compared to adults. Their return-sweeps were launched closer to the end of the line and landed closer to the left margin. This indicates that developing readers must foveate more of the text to aid reading comprehension. Similar to adults, children's line-final and undersweep-fixations were shorter than their intra-line fixations. However, children's accurate line-initial fixations were also shorter than their intra-line fixations contrary to adult's data. We say more about this below.

#### **4.1 Return-sweep and corrective saccade parameters.**

Adults were able to launch their return-sweeps from a location further from the end of the line than did children. It would, therefore, seem that adults do not need to fixate the extremes of lines in order to encode the text there. The effect on landing site is of particular interest because the target of the return-sweep lies far outside the parafovea. That is, even adults would be unable to encode the letters at the start of a new line during the final fixation on the prior line. Instead, adults seem to have learned to target further into lines due to the parafoveal encoding they will be able to do at the start of a new line following a return-sweep. These findings highlight how skilled reading relies on parafoveal processing. Children's preference for viewing positions closer to the left margin is further reflected by the increased prevalence of corrective saccades despite their line initial landing positions

being closer to the left margin than adults. This again suggests that children rely more on foveal processing to encode information.

#### **4.2 Return-sweep fixation durations.**

In comparison to adults, children are slower at reading and exhibit more frequent and longer fixations (Blythe, 2014; Blythe & Joseph, 2011; Reichle et al., 2013). We replicated longer intra-line reading fixations for children. Additionally, we report that line-final fixations are shorter than intra-line fixations for both adults and children. We see two possible explanations for this effect. The first is consistent with Rayner (1977) and suggests that adults and children equally benefit from reduced parafoveal processing demands as they approach the end of a line. However, given that children in the current study are likely to rely on foveal processing (cf. Häikiö et al., 2009), such an account would predict a smaller reduction in line-final fixation durations in children. This was not the case. The alternative is consistent with Abrams and Zuber (1972) and postulates that a reduction in fixation duration for line-final fixations is the result of reduced lexical processing in the face of return-sweep planning. By this interpretation, a similar slope between groups indicates that skilled and developing reader's time course of return-sweep planning is similar. This seems plausible given similarities in general intra-line saccade targeting. However, to draw a firm conclusion further research is required to assess the extent to which lexical information modulates line-final fixation durations in adults and children.

For adults, accurate line-initial fixations were longer than intra-line fixations. However, children's accurate line-initial fixations did not differ from their intra-line fixations. We explain this finding with regards to the differences in parafoveal processing between adults and children. When readers land at the start of the line, they must build a visual representation of the first few characters on a line (Kuperman et al., 2010). By landing further into the line adult readers will have more information to the left of fixation to add to

their ongoing representation than children. Assuming that a cost of parafoveal processing is increased fixation durations, this would subsequently increase line-initial fixation durations for adults. In contrast, children land closer to the left margin in accurate cases and would require less parafoveal processing to process information to the left of fixation. This may result in relatively reduced line initial fixations for younger readers.

In addition to processing information to the left of fixation, children in our study will have processed less information to the right of fixation. This is because children made shorter forwards saccades following an accurate return-sweep than adults. This indicates that adults will have encoded more information parafoveally prior to the saccade following an accurate return-sweep. This increased parafoveal processing may have resulted in longer accurate line-initial fixation durations for adults compared to their intra-line fixations. Interestingly, adults intra-line saccades were shorter than those following an accurate return-sweep. This suggests that adult's intra-line fixations may be shorter than accurate line-initial fixations due to differential parafoveal processing strategies across the line of text (e.g. start-up effects; c.f. Kuperman et al., 2010).

Undersweep-fixation durations were the shortest of all fixation populations, with durations not differing between groups. A lack of difference in duration despite the clear differences in reading ability demonstrates that these fixations are under oculomotor control rather than lexical control. This provides further support for the empirical distinction between accurate line-initial fixations and undersweep-fixations. With regards to differences between adult's and children's reading, this suggests that children are just as efficient in responding to the retinal error signal that results from landing in an unintended location as adults. Furthermore, differences in attentional and linguistic processing appear responsible for the differences observed in adult's and children's intra-line reading fixation time measures.

## 5. Conclusions

We report a novel study in which the return-sweep and corrective saccade parameters were reported for adults and children reading multiline texts. We observed several differences in adult's and children's return-sweep behaviour. Children tended to launch their return-sweeps closer to the end of a line than did adults, while their return-sweeps landed closer to the left margin. Even though children landed closer to the start of line than adults, they still initiated corrective saccades more frequently. This shows that, compared to adults, children adopt a reading strategy in which they fixate more extreme locations on a line to compensate for reduced parafoveal processing.

The findings from the current study also provide several benchmarks for computational models of eye movement control, such as E-Z Reader (Reichle & Sheridan, 2015) and SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005). As currently implemented, these models do not make predictions about return-sweep behaviour. In order to model the reading of books such as Harry Potter, these models of eye movement control will have to consider fixations either side of the return-sweep and the factors contributing to return-sweep error. To develop more accurate models, additional research will be needed to explore how visual, attentional, and linguistic processes are influenced by return-sweeps.

### Notes

1. The stimuli had been developed for children aged 7 and above. However, parents of children aged 6 had returned consent forms and these children wished to take part. Rather than refuse these children, we collected data and report it in the main analysis. All children in the current study had a TOWRE reading age that was greater than 7-years old. When identical analyses were conducted for children aged 7 and above, the patterns were consistent with what is reported in the main body.
2. Two items had words with an AoA above 7.00 years. Additional analysis which excluded these items did not change the overall pattern of results.
3. An LME model fit to log-transformed undersweep-fixation duration indicated that undersweep-fixation durations were comparable between adults and children,  $b = -.02$ ,  $SE = .01$ ,  $t = -1.66$ ,  $p = .101$ .

### **Acknowledgments**

This data was presented at the 59<sup>th</sup> Annual Meeting of the Psychonomic Society in New Orleans, Louisiana, USA. We would like to thank Amy Burbage, Millie Young, and Meltem Erenus for their assistance with data collection. In addition, we thank the parents and children who took part in this study as unpaid volunteers. We thank Hazel Blythe and an anonymous reviewer for their helpful comments on a prior version.

**Contributors**

Conceived and designed the study: AJP, TJS, JAK. Performed the study: AJP. Analysed the data: AJP. Wrote the paper: AJP, TJS, JAK.

**Declarations of interest**

None.

**Role of funding source**

This work was made possible in part by a research grant from the Microsoft Corporation to TJS which was matched by Bournemouth University to fund AJP's doctoral studies.

### References

- Abrams, S. G., & Zuber, B. L. (1972). Some temporal characteristics of information processing during reading. *Reading Research Quarterly*, 8, 40-51.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, 59, 390-412.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Bates et al. (2018). lme4: Linear Mixed-Effects Models using 'Eigen' and S4. R package version 1.1-17.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967.
- Becker, W. (1976). Do corrective saccades depend exclusively on retinal feedback? A note on the possible role of non-retinal feedback. *Vision Research*, 41, 425-427.
- Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science*, 23, 201–207.
- Blythe, H. I., & Joseph, H. S. S. I. (2011). Children's eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 643–662). Oxford: Oxford University Press.
- Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision research*, 51, 84-92.
- Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research*, 49, 1583–1591.

- Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular coordination of eye movements during reading in children and adults. *Vision Research*, *46*, 3898–3908.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977-990.
- Bucci, M. P., & Seassau, M. (2012). Saccadic eye movements in children: a developmental study. *Experimental brain research*, *222*, 21-30.
- Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: effects of reading skill. *Canadian Journal of Experimental Psychology*, *59*, 209-217.
- Doré-Mazars, K., Vergilino-Perez, D., Lemoine, C., & Bucci, M. P. (2011). Adaptation of reactive saccades in normal children. *Investigative Ophthalmology & Visual Science*, *52*, 4813-4818.
- Engbert, R., & Kliegl, R. (2011). Parallel graded attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 787-800). New York, NY, US: Oxford University Press.
- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813.
- Häikiö, T., Bertram, R., & Hyönä, J. (2010). Development of parafoveal processing within and across words in reading: Evidence from the boundary paradigm. *The Quarterly Journal of Experimental Psychology*, *63*, 1982-1998.
- 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*, 167-181.

- Hawley, T. T., Stern, J. A., & Chen, S. C. (1974). Computer analysis of eye movements during reading. *Literacy Research and Instruction, 13*, 307-317.
- Heller, D. (1982). Eye movements in reading. In R. Groner, & P. Fraisse (Eds.), *Cognition and Eye Movements* (pp. 487-498). Berlin, Deutscher Verlag der Wissenschaften.
- Heller, D., & Radach, R. (1992). Returning to an unanswered question: On the role of corrective saccades in text reading [Abstract]. *International Journal of Psychology, 27*, 55.
- Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 417-429.
- Hofmeister, J. (1997). On corrective saccades in reading and reading-like tasks. Unpublished doctoral dissertation. Rheinisch-Westfälische, Aachen
- Hofmeister, J., Heller, D., & Radach, R. (1999). The return sweep in reading. In W. Becker, H. Deubel, & T. Mergner (Eds.), *Current Oculomotor Research* (pp. 349-357). United States: Springer.
- Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision research, 49*, 2948-2959.
- Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1430-1440.
- Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research, 49*, 2078-2086.

- Joseph, H. S., Nation, K., & Livversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review, 42*, 207-222.
- Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye-movements in sentence and paragraph reading. *The Quarterly Journal of Experimental Psychology, 63*, 1838–1857.
- Kuperman, V., Stadthagen-Gonzalez, H. & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behaviour Research Methods, 44*, 978-990.
- Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2016). lmerTest: Tests in linear mixed effects models. R package version 2.0-30.
- Luna, B., Velanova, K., & Geier, C. F. (2008). Development of eye-movement control. *Brain and cognition, 68*, 293-308
- Marx, C., Hawelka, S., Schuster, S., & Hutzler, F. (2015). An incremental boundary study on parafoveal preprocessing in children reading aloud: Parafoveal masks overestimate the preview benefit. *Journal of Cognitive Psychology, 27*, 549-561.
- Marx, C., Hawelka, S., Schuster, S., & Hutzler, F. (2017). Foveal processing difficulty does not affect parafoveal preprocessing in young readers. *Scientific Reports, 7*, 41602.
- Marx, C., Hutzler, F., Schuster, S., & Hawelka, S. (2016). On the development of parafoveal preprocessing: Evidence from the incremental boundary paradigm. *Frontiers in psychology, 7*, 514.
- 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*, 1107-1118.
- McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. *Vision and visual dyslexia, 13*, 251-262.

- 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*, 266-293
- Netchine, S., Guihou, M. C., Greenbaum, C., & Englander, G. (1983). Retour a la ligne, age des lecteurs et accessibilite au texte. *Le Travail Humain*, *46*, 139-153.
- O'Regan, J. K., & Jacobs, A. M. (1992). Optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 185.
- O'Regan, J. K., Lévy-Schoen, A., Pynte, J., & Brugailière, B. É. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 250.
- Pagán, A., Blythe, H. I., & Liversedge, S. P. (2016). Parafoveal preprocessing of word initial trigrams during reading in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 411.
- 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*, 902-911.
- Pollatsek, A., Juhasz, B. J., Reichle, E. D., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: A reversed delayed effect of word length. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 726.
- Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 148.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

- Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65-81.
- Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition*, 4, 443-448.
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of experimental child psychology*, 41, 211-236.
- Rayner, K. (1998). Eye movements in reading in formation processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62, 1457-1506.
- Rayner, K., Liversedge, S. P., White, S. J., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye movements. *Psychological science*, 14, 385-388.
- Rayner, K., White, S. J., & Liversedge, S. P. (2006). Raeding wrods with jubmled lettres: There is a cost. *Psychological Science*, 17, 192-193.
- Reichle, E. D. (2011). Serial-attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 767-786). New York, NY, US: Oxford University Press.
- 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. 277-290). Oxford: Oxford University Press.
- Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L, 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, 110-149.

- Salman, M. S., Sharpe, J. A., Eizenman, M., Lillakas, L., Westall, C., To, T., Dennis, M., & Steinbach, M. J. (2006). Saccades in children. *Vision research*, *46*(8-9), 1432-1439.
- Schroeder, S., Hyönä, J., & Liversedge, S. P. (2015). Developmental eye-tracking research in reading: Introduction to the special issue. *Journal of Cognitive Psychology*, *27*, 500-510.
- Shebilske, W.L., 1975. 'Reading eye movements from an information-processing point of view'. In: D.W. Massaro (ed.), *Understanding language: An information-processing analysis of speech perception, reading, and psycholinguistics*. New York: Academic Press.
- Stern, J. A. (1981). Eye movements, reading and cognition. In J. W. Senders, D.F. Fisher, & R. A. Monty (Eds.), *Eye movements and higher psychological functions* (pp. 145-155). New York: Erlbaum.
- Tiffin-Richards, S. P., & Schroeder, S. (2015). Children's and adults' parafoveal processes in German: Phonological and orthographic effects. *Journal of Cognitive Psychology*, *27*, 531-548.
- Tiffin-Richards, S. P., & Schroeder, S. (2018). The development of wrap-up processes in text reading: A study of children's eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1051-1063.
- Tinker, M. A. (1963). *Legibility of print*. Ames: Iowa State University Press.
- Torgesen, J. K., Rashotte, C. A., & Wagner, R. K. (1999). *TOWRE: Test of word reading efficiency*. Austin, TX: Pro-ed.
- Trauzettel-Klosinski, S., Koitzsch, A. M., Dürrwächter, U., Sokolov, A. N., Reinhard, J., & Klosinski, G. (2010). Eye movement in German-speaking children with and without dyslexia when reading aloud. *Acta Ophthalmologica*, *88*, 681-691.

- van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*, 1176-1190.
- Wolf, M., & Denckla, M. B. (2005). *RAN/RAS: Rapid automatized naming and rapid alternating stimulus tests*. Austin, TX: Pro-ed.