Reading disappearing text: Why do children refixate words?

Hazel I. Blythe[1]*, Tuomo Häikiö[2], Raymond Bertam[2], Simon P. Liversedge[1] & Jukka Hyönä[2]

[1]School of Psychology, University of Southampton, UK

[2]Department of Psychology, University of Turku, Finland

*Corresponding author

School of Psychology

Shackleton Building

University of Southampton

Highfield

Southampton, SO17 1BJ

United Kingdom

Email: hib@soton.ac.uk

Tel: +44 23 8059 2917

Fax: +44 23 8059 4597

Word count = 7,990

Key words: children; eye movements; reading; word length

This is the final, accepted version of the article published in
Vision Research, 51(1), 84-92. https://doi.org10.1016/j.visres.2010.10.003

Abstract

We compared Finnish adults' and children's eye movements on long (8-letter) and short (4-letter) target words embedded in sentences, presented either normally or as disappearing text. When reading disappearing text, where refixations did not provide new information, the 8- to 9-year old children made fewer refixations but more regressions back to long words compared to when reading normal text. This difference was not observed in the adults or 10- to 11-year old children. We conclude that the younger children required a second visual sample on the long words, and they adapted their eye movement behaviour when reading disappearing text accordingly.

1. Introduction

When reading, we occasionally make two fixations on a word before moving our eyes to a different word in the sentence. The role that these refixations play in reading is unclear. Intriguingly, the data from some studies using the disappearing text paradigm have suggested that we do not necessarily need to make refixations when we read; that refixations are, to some degree, redundant (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Liversedge, Rayner, White, Vergilino-Perez, Findlay & Kentridge, 2004; Rayner, Liversedge, White & Vergilino-Perez, 2003; Rayner, Liversedge & White, 2006; Rayner, Yang, Castelhano & Liversedge, 2010).

In the disappearing text paradigm, sentences are presented in a gaze-contingent display with invisible boundaries between all words in the sentence. Whenever the reader's eye crosses a boundary between words, a timer is set to count down to a pre-specified delay (typically, 60 ms). Following this delay, the newly fixated word disappears, so that the participant is fixating a blank space. Once they move their eyes to fixate a different word in the sentence, the previously fixated word reappears and the newly fixated word disappears after the specified delay. Thus, there is only one word missing from the sentence at any time, but it is the word that the reader is fixating. In this way, the reader's opportunity to visually sample the fixated word is temporally limited to the initial portion of the fixation on that word (as well as any parafoveal pre-processing that had occurred).

When reading disappearing text, there was no overall cost to participants' sentence reading times, nor was comprehension impaired. There was a trade-off between decreased refixation probability and increased fixation durations, such that sentence processing times were equal (Liversedge *et al.*, 2004; Rayner *et al.*, 2003, 2006). However, this trade-off suggests that refixations are not essential for visually re-sampling the fixated word.

During normal text reading, refixations offer the reader a second opportunity to sample the visual information from that word. The younger the reader, the more frequently refixations are made; Blythe *et al.* (2009) found that 15-25% of words were refixated by 7- to 9-year old children. When the text has disappeared during the initial fixation on a word, there is no opportunity to usefully refixate that word. It might, therefore, be expected that a compensatory strategy would be employed, such as increasing the number of regressions made, in order to obtain a second fixation on that word and so, to successfully read disappearing text. This, however, was not found to be the case; readers did not make more regressions when reading disappearing text (Blythe *et al.*, 2009; Liversedge *et al.*, 2004; Rayner *et al.*, 2003, 2006; Rayner *et al.*, 2010). Furthermore, effects of word frequency on reading times on 6-letter target words occurred under disappearing text conditions, for the youngest children tested (7-years) as well as adults, showing that even very brief visual inputs of 40-57 ms were sufficient for readers to initiate normal lexical processing. It seems that, despite the typically high frequency with which younger children (7- to 9-years) make a second fixation on a word, they do not necessarily require this second fixation for successful reading.

An important point to note is that the target words in the Blythe *et al.* study were six letters long. Another recent study which has examined children's eye movements during reading found significant effects of word length upon refixation probability, such that the longer the word, the more likely a reader was to refixate that word (Joseph, Liversedge, Blythe, White & Rayner, 2009; see also Hyönä & Olson, 1995). Furthermore, this effect was stronger in children aged 7- to 11-years than in adults.

These data suggest that younger readers can encode less visual information within a fixation compared to older readers and so may require more fixations on a word in order for lexical identification to occur. This is based on the idea that children's encoding of visual information during fixations is somehow slower or less fluent than that of adults. As beginning readers, children are inherently less familiar with the printed forms of words than are adults. Thus, it may be the case that children require more visual samples of a word in order to encode it sufficiently for the initiation of normal lexical processing, and so a second fixation may often be made in order for successful lexical identification to occur. This argument is consistent with previous research showing that the perceptual span is smaller in children compared to adults (Häikiö *et al.*, 2009; Rayner, 1986). Importantly, in the Joseph *et al.* study the word length manipulation compared 4-letter and 8-letter words.

It seems likely that the discrepancy in results between Blythe *et al.* and Joseph *et al.* is a consequence of the length of target words that were used in each of these studies. To be clear, it may well be the case that refixations provide a necessary second opportunity for beginning readers to visually sample 8-letter words, but this second visual sample is not required for 6-letter words since they are shorter.

The present experiment was designed to test the hypothesis that refixations serve the purpose of allowing a second visual sample on long words (8 letters or more) which allows lexical processing to proceed unhindered, which is not necessary for short words (6 letters or less). We used the disappearing text paradigm to compare reading of 8-letter and 4-letter words in Finnish, a language where long words are relatively common (compared to English). Word length is an important aspect of the Finnish language. Many compound words are concatenated to form one long word (rather than the compound word being spaced or hyphenated which is more common in English) (see Hyönä, Bertram & Pollatsek, 2004; Juhasz, 2008).

In English, long words tend to have relatively low frequencies and so any manipulation that controls for frequency will result in stimuli that include long and short words that are relatively uncommon. Due to the relatively high occurrence of long words in Finnish, children in Finland know more long words than do English children of a comparable age, and those words are more familiar to them. Thus, it is possible to make a far stronger word length manipulation for Finnish children (keeping word frequency relatively high so that the children can easily read the target words) than can be done for English children. This was another principle reason why we carried out this study with Finnish child participants.

To reiterate, there is an apparent contradiction in the literature on children's eye movements during reading. While refixations on 6-letter words appear to be unnecessary (Blythe *et al.*, 2009), words of 8-letters do elicit refixations proportionally more often in children than in adults (Joseph *et al.*, 2009). Thus, 8-letter and longer words seem to be the cases where a second visual sample may be necessary for word identification to proceed unhindered. We tested this hypothesis in Finnish, where word length is a particularly important aspect of the language. We measured the eye movements of children aged 8- to 9- years and 10- to 11-years, as well as a group of skilled adult readers, as they read sentences that contained a target word that was manipulated for length. These sentences were presented either normally, or as disappearing text where each word disappeared in turn 60 ms after fixation onset on that word.

In addition to the discrepancy in the literature with respect to children's refixations and word length, our study was also relevant to several other, broader theoretical issues with respect to eye movement control during reading. First, we examined age-related differences in eye movement behaviour during reading by comparing older (10- to 11-years) and younger children (8- to 9-years) as they read. While there is a vast literature on the eye movements of skilled adult readers (see Rayner, 1998, 2009 for reviews), very little eye movement research has been conducted to examine children's reading (see Blythe & Joseph, 2010, for a review). While recent work has begun to address this imbalance, it is still the case that eye movement researchers have a greater understanding of the "end state" with respect to reading than of the development of this process.

More specifically, this experiment allowed us to consider the present data from Finnish children in relation to those from English children of comparable age using the same experimental paradigm reported by Blythe *et al.* (2009). This offers the opportunity to consider how language-specific aspects of printed text (such as the relative frequency of long words) impact on age-related changes in eye movement behaviour during reading. While many researchers have used eye movements to examine reading in languages other than English, both for adults and children with dyslexia (for example, Finnish adults, e.g., Hyönä *et el.*, 2004; Chinese, American, Japanese, and Korean adults, e.g., Shen *et al.*, 2010; Italian and German children with dyslexia, see Kirkby, Webster, Blythe, & Liversedge, 2008, for a review), only one study has examined this issue in typically developing beginning readers by discussing differences in eye movement behaviour between English- and Finnish-speaking children who were tested on the same experimental paradigm (Häikiö *et al.*, 2009). Thus, this experiment allowed us the opportunity to examine age-related changes in eye movement behaviour during reading, and to consider the differences in eye movement control during reading between Finnish and English children.

There were three alternative outcomes with respect to our question of whether children require a second visual sample on long words. First, that a second visual sample is required on longer words and that, consequently, the disappearing text manipulation would cause disruption to reading for the 8-letter words but not the 4-letter words. This disruption would be reflected by increased total fixation times (the sum of all fixations on the target word) for longer words and also, presumably, increased regression frequencies back onto the target word in order that participants might gain the required second visual sample through an alternative strategy. Furthermore, on the basis of the literature showing that children have a smaller perceptual span than adults (Häikiö *et al.*, 2009; Rayner, 1986), it might also be expected that the disruption associated with long words would be more pronounced in children compared to adults.

Alternatively, it was possible that a second visual sample is required on longer words, but that the disappearing text manipulation would not cause disruption to reading in terms of increased reading times. That is, it might be the case that there is an increase in regressions back onto the target word in order to obtain a second visual sample, but that this alternative strategy is so effective that it would not lead to increased overall reading times. This possibility would be supported if we were to find shorter gaze durations (the sum of all fixations on the target word in first-pass reading) for the disappearing text condition in comparison to the normal text condition, but equal total fixation times for both conditions.

The third possibility was that the typically observed refixations in children's reading serve some purpose other than providing a necessary, second visual sample, in which case there should be no more disruption for long words than for short words when presented as disappearing text. Here, we would expect to see the previously documented, strategic differences between disappearing text and normal text - fewer refixations but longer fixation durations and no increase in regressions back onto the target word, with no overall cost to total fixation times for both long and short words (Blythe *et al.*, 2009; Liversedge *et al.*, 2004; Rayner *et al.*, 2003). In other words, we would expect to see exactly the same pattern of eye movement behaviour elicited by the disappearing text manipulation for long and short words.

2. Method

2.1. Participants. All participants were native Finnish speakers with normal or corrected-to-normal vision and no known reading difficulties. The participants participated voluntarily in the study and all participants were naïve regarding the purpose of the study. Adult participants were paid for their participation. Child participants were unpaid volunteers, but were given candy by way of thanks for their participation. The 16 adult participants were all students at the University of Turku, with an age range of 19- to 27-years. The 32 child participants were volunteers from a local school in the Turku area. The 16 younger children were aged 8- to 9-years, and the 16 older children were aged 10- to 11-years (2nd and 4th grade, respectively).

2.2. Apparatus. Monocular eye movement recordings from the right eye were taken using an EyeLink 1000 eye tracker. The position of participants' right eye was recorded every millisecond. The eye tracker was interfaced with a Dell Optiplex GX745 computer, with all sentences presented on a 20 inch ViewSonic G225f monitor that was set at a refresh rate of 120 Hz. Sentences were presented in black, Courier New font size 14, on a white background. Sentences were presented at a viewing distance of 60 cm. Participants were asked to lean on a chin cup and against forehead rests during the experiment, to eliminate head movements.

2.3. Materials and Design. Sixty experimental sentences were constructed, each containing a 4-letter (short) or 8-letter (long) monomorphemic target word. Word frequencies were derived from the unpublished computerized Turun Sanomat newspaper corpus with the help of the WordMill database program of Laine and Virtanen (1999). The mean frequency of short words was 200 counts per million, and the mean frequency of long words was 201 counts per million; there was no significant difference in frequency between the long and short words (t_2 (29) = 0.07, p = 0.85). For each pair of target words, the sentence frames were identical and neutral in that they did not constrain the identity of the target word; after that, the sentences differed. This allowed all participants to read all 60 sentence frames, without noticing any repetition. All sentences were between 35 and 74 characters long (between 5 and 13 words), extending from the left side of the screen. An example of the stimuli with the word length manipulation is given in Table[1]. In addition to the 60 experimental items, six practise items were presented at the beginning of each block. After 12 of the sentences, participants were required to respond to simple comprehension statements in relation to the experimental sentences, making a "true/ false" response using a button box. The sentences were specifically written so that the children would have no difficulty understanding them.

Condition	Sentence
Long	Tien päässä näkyi **sairaala** ja sen takana urheilukenttä. *There was a **hospital** at the end of the road, and a sports field behind it.*
Short	Tien päässä näkyi **halli** ja sen vieressä parkkipaikka. *There was a **hall** at the end of the road, and a parking lot next to it.*

Table 1. Sample experimental sentences containing a target word (bold, underlined) that was either long (8 letters) or short (4 letters). The English translation is given in italics underneath each sentence.

Half of the sentences were presented normally and half of the sentences were presented under disappearing text conditions; text presentation condition was blocked within the experiment and presentation of these blocks was counterbalanced across participants. In the disappearing text condition, each word disappeared 60 ms after fixation onset on that word. Once the reader moved their eyes onto a new word in the sentence, the previously fixated word reappeared but the newly fixated word then disappeared 60 ms after fixation onset. Note that using identical experimental conditions Rayner *et al.*, (2003) demonstrated that there was no afterimage on the screen due to phosphor persistence that would have allowed participants to perceive the words after they had disappeared.

2.4 Procedure. Participants were instructed to read the sentences normally, and to answer the questions as accurately as possible by pressing a game pad to indicate "yes/no" responses. All instructions were given verbally to the children both at the start and throughout the experiment as required. Children were given lots of encouragement throughout the experiment.

Viewing was binocular, but only the right eye was recorded. An initial calibration of the eye tracker was carried out in which the participant was instructed to look at grid of nine fixation points while their fixation position was recorded for each point. Once the eye-tracker had been calibrated with satisfactory accuracy, the sentences were presented. Following every sentence, the calibration was checked for accuracy, and the eye tracker was recalibrated whenever necessary. All participants were given a break half way through the experiment, and additional breaks were given whenever required. The entire experiment lasted approximately 20 minutes for adults, and half an hour for children.

2.5. Analyses. Data were only analysed from participants who had scored at least 70% on the comprehension questions, in order to ensure that they had understood the sentences. Comprehension scores, refixation probability, and word skipping probability were analysed using loglinear regressions. For reading time measures and the numbers of fixations and regressions made per sentence, 3 (age group: adults, 10- to 11-years, and 8- to 9-years) x 2 (presentation condition: normal or disappearing text) x 2 (word length: 4 or 8 letters) analyses of variance (ANOVAs) and t-tests were conducted treating participants (*F*, *t*) as a random variable (Raaijmakers, Schrijnemakers, & Gremmen, 1999). Consistent with Raaijmakers *et al.*, our target words were matched for frequency across the word length manipulation. Word length was the only between-items manipulation in our experimental design. The items were divided between two counterbalanced lists (between-participants) so that all items were seen in both the normal and the disappearing text condition; hence, there was no need to match the items for frequency across text presentation conditions. Given this matching of items and counterbalancing, we report F_1 but not F_2 analyses (Raaijmakers *et al.*, 1999).

T-tests were conducted in order to examine reliable main effects of participant group and interaction terms from the ANOVAs in more detail. A Bonferroni correction was applied to these t-tests in order to reduce the likelihood of making a Type I error. For significant main effects of group in all analyses, three t-tests were conducted to compare the three participant groups with each other and so a Bonferroni-corrected *p*-value of less than 0.02 was accepted as significant. Similarly, for the group by condition interaction term in the global analyses, three t-tests were conducted to compare the effect of the disappearing text manipulation in each participant group separately, and so a Bonferroni-corrected *p*-value of less than 0.02 was accepted as significant. In the local analyses, for the group by length interaction term, two t-tests were conducted to examine the differences in each participant group separately and so, again, a Bonferroni-corrected *p*-value of less than 0.03 was accepted as significant. For the condition by length interaction term, two t-tests were conducted in order to examine the effect of text presentation condition separately for each word length, and so a Bonferroni-corrected *p*-value of less than 0.03 was accepted as significant. Finally, for the three-way interaction term, six t-tests were conducted to examine the effect of text presentation condition separately for each group and each word length, and so a Bonferroni-corrected *p*-value of less than 0.01 was accepted as significant.

3. Results

3.1. Response accuracy. We analysed response accuracy to the comprehension questions in order to confirm that participants of all ages were able to successfully read and comprehend the sentences under both normal and disappearing text conditions (mean scores across conditions are shown in Table 2). There were no significant differences between the participant groups (*Z*s < 1.5, *p*s > 0.1), there was no significant effect of text presentation condition ($Z = 0.54$, $SE = 0.54$, $p = 0.59$), and there were no significant interactions (*Z*s < 1, *p*s > 0.4). Thus, participants were able to read and comprehend the sentences equally well under both normal and disappearing text presentation conditions.

3.2. Global analyses. These analyses are based on all fixations across the sentences, in both normal and disappearing text conditions. There was no detrimental effect of the disappearing text manipulation on total sentence reading times; the main effect of text presentation did not approach significance (F (2, 45) = 0.48, p = 0.49). Younger readers had longer sentence reading times overall compared to older readers (F (1, 45) = 26.47, p < 0.001), and t-tests showed that all three groups differed significantly from each other (all ts > 3, all ps ≤ 0.001). However, the interaction between group and text presentation was not reliable (F (2, 45) = 0.68, p = 0.51). More detailed analyses will now be presented considering the different patterns of eye movements elicited by the two text presentation conditions for readers of different ages.

The pattern of results for fixation durations, the number of fixations per sentence, refixation probability, the number of regressions per sentence, and word skipping probability was similar across these five measures. First, there were main effects of participant group on all measures – the 8- to 9-year-olds had the longest fixation durations, made the most fixations regressions per sentence while the adults had the shortest fixation durations and made the fewest fixations and regressions per sentence (all Fs > 6, all ps < 0.01). On all these measures, all three participant groups differed significantly from each other (all ts > 2, all ps ≤ 0.02). The only exception was the comparison of older and younger children for the number of regressions per sentence (t (30) = 0.93, p = 0.36). There were also effects of age group on both refixation probability and word skipping probability. Both adults (Z = 21.43, SE = 0.06, p < 0.001) and older children (Z = 12.37, SE = 0.05, p < 0.001) made fewer refixations than younger children, and skipped more words (adults: Z = 12.38, SE = 0.07, p < 0.001; older children: Z = 8.92, SE = 0.07, p < 0.001). These differences between groups, therefore, clearly reflect the changes in eye movement behaviour that characterise the progression from beginning to skilled reader.

Second, on all measures there was a significant effect of text presentation. When reading disappearing text, readers had longer fixation durations (F (2, 45) = 24.47, p < 0.001), made fewer fixations (F (2, 45) = 13.56, p = 0.001), but more regressions per sentence (F_1 (2, 45) = 9.75, p < 0.01), and made fewer refixations (Z = 26.15, SE = 0.06, p < 0.001) but skipped more words (Z = 4.11, SE = 0.08, p < 0.001) compared to when reading normally presented text.

	Adults		10- to 11-years		8- to 9-years	
	Disappearing text	Normal text	Disappearing text	Normal text	Disappearing text	Normal text
Comprehension score	98%	99%	93%	95%	94%	92%
Total sentence reading time	2014 (918)	2092 (839)	2973 (1075)	3213 (1042)	4642 (2503)	4563 (1860)
Fixation duration	210 (107)	206 (103)	238 (125)	232 (128)	294 (189)	253 (149)
Number of fixations per sentence	9.6 (3.5)	10.1 (3.4)	12.5 (4.3)	13.9 (4.0)	15.8 (7.6)	18.0 (7.3)
Number of regressions per sentence	1.9 (1.6)	1.5 (1.5)	3.2 (2.0)	2.9 (1.8)	4.1 (3.4)	3.1 (2.6)
Refixation probability	10% (7%)	19% (10%	8% (4%)	30% (10%)	14% (7%)	45% (15%)
Word skipping probability	26% (8%)	20% (7%)	19% (7%)	16% (7%)	12% (7%)	9% (6%)

Table 2. Means and standard deviations for total sentence reading times, fixation durations, number of fixations per sentence, number of regressions per sentence, refixation probability and word skipping probability for the three participant groups under both normal and disappearing text conditions Standard deviations are given in parentheses.

There were significant interactions between participant group and text presentation condition for mean fixation duration and refixation probability. As can be seen in Table 2, for mean fixation duration (F (2, 45) = 11.03, p < 0.001), the effect was largest and significant for the younger children (t (15) = 4.72, p < 0.001), was smaller and marginal for the older children (t (15) = 1.87, p = 0.08), and was smallest and non-significant for the adults (t (15) = 0.71, p = 0.49). For refixation probability, again the effect size decreased with age, being equal for older and younger children (Z = 0.66, SE = 0.10, p = 0.51) but significantly smaller for adults (Z = 8.41, SE = 0.10, p < 0.001). The interaction between age group and text presentation condition was not significant for the number of regressions made per sentence (F (2, 45) = 1.84, p = 0.17), the number of fixations made per sentence (F (2, 45) = 1.67, p = 0.20), or word skipping probability (Zs < 2, ps > 0.1).

In summary, all participants showed significant differences in their eye movement behaviour when reading disappearing text compared to when reading normal text. Participants had longer fixation durations, made fewer fixations but more regressions per sentence, had a lower refixation probability, and skipped more words when reading disappearing text compared to normal text although there was no overall cost to sentence reading times. Further, the difference between conditions was greatest for 8- to 9-year olds; this interaction was significant for fixation durations and refixation probability.

3.3. Local analyses.

3.3.1. Target words. These analyses are based on the subset of data for the target word in each sentence that was manipulated for length, being either 4 or 8 letters long.

		First fixation duration (ms)		Gaze duration (ms)		Total fixation time (ms)	
		Long	Short	Long	Short	Long	Short
Adults	Normal text	201 (57)	209 (59)	238 (90)	223 (69)	282 (140)	256 (139)
	Disappearing text	190 (69)	190 (60)	206 (81)	200 (69)	265 (194)	212 (200)
	Normal	235	245	342	292	478	402

10- to 11- years	text	(77)	(90)	(176)	(120)	(278)	(232)
	Disappearing text	242 (59)	232 (81)	249 (69)	240 (92)	404 (259)	371 (197)
8- to 9- years	Normal text	281 (141)	266 (110)	591 (425)	381 (190)	757 (520)	545 (331)
	Disappearing text	318 (177)	277 (112)	385 (305)	311 (179)	772 (918)	528 (434)

Table 3. Means and standard deviations for first fixation durations, gaze durations, and total fixation times on long (8 letter) and short (4 letter) words. Standard deviations are given in parentheses.

First, we examined reading times on these target words. Means and standard deviations for first fixation durations, gaze durations and total fixation times on long and short words for all three participant groups under both text presentation conditions are shown in Table 3. The main aim was to assess whether processing of long words was affected by text presentation to the same extent as processing of short words.

First fixation duration. There was no effect of text presentation (F (1, 45) = 0.12, p = 0.74), but a marginal effect of word length on first fixation durations (F (1, 45) = 3.64, p = 0.06). There was also a significant interaction between participant group and word length for first fixation duration (F (2, 45) = 4.53, p = 0.02). The younger children showed a marginal effect of word length on first fixation durations (t (15) = 2.41, p = 0.03), while there was no effect of word length for the older children or the adults (both ts < 2, all ps > 0.2). There was also a significant interaction between word length and text presentation condition (F (2, 45) = 6.14, p = 0.02). While the difference in first fixation duration between normal and disappearing text was not significant for either long or short words (both ts < 2, both ps > 0.1), the numerical difference between text presentation conditions was greater for long words (11 ms) than for short words (4 ms). The three-way interaction between group, text presentation condition, and word length was not significant (F (2, 45) = 0.46, p = 0.64).

Gaze duration. On gaze durations, there was a significant main effect of text presentation (F (1, 45) = 44.52, p < 0.001). Target words elicited shorter gaze durations when presented in disappearing text format than when presented in normal text format. Additionally, there was a significant main effect of word length (F (1, 45) = 31.04, p < 0.001), showing that long words elicited longer gaze durations than short words. The effect of text presentation condition was qualified by an interaction with word length (F (2, 45) = 14.77, p < 0.001). Gaze durations were shorter in the disappearing text condition compared to the normal condition for both long and short target words (both ts > 5, both ps < 0.001), but the effect was numerically greater for long words (109 ms) than for short words (46 ms). These data are shown in Panel A of Figure 1, alongside data giving mean total fixation times, and refixation and regression probabilities for the long and short target words.

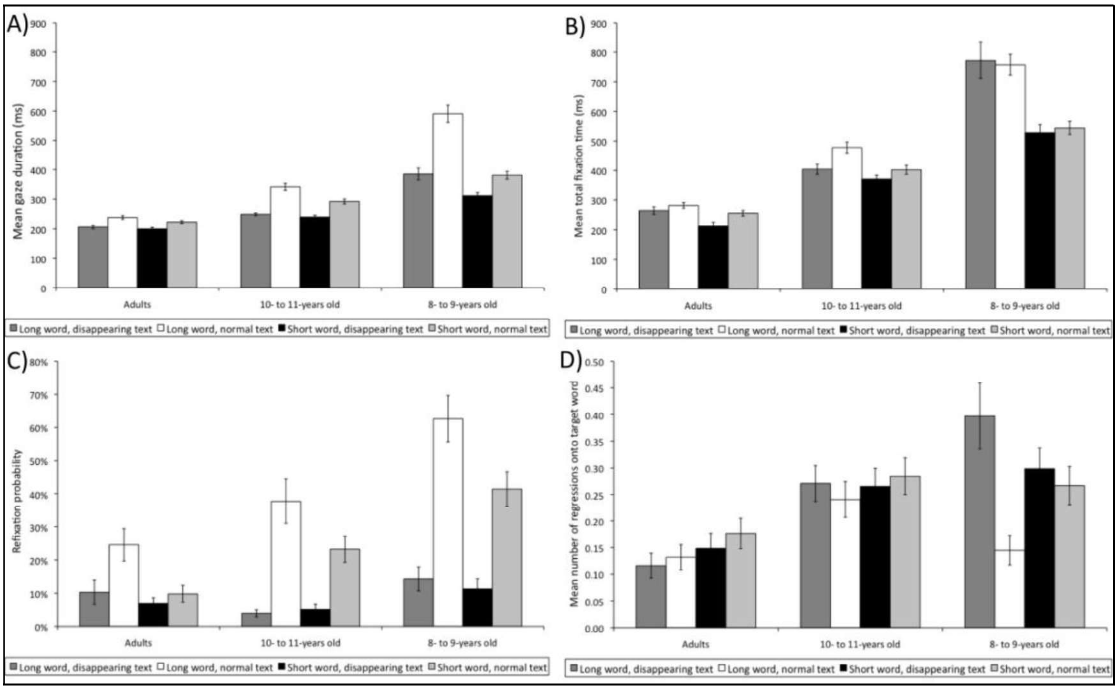


Figure 1. Mean gaze duration (Panel A), mean total fixation time (Panel B), refixation probability (Panel C), and mean number of regressions in (Panel D) for long and short target words for the three participant groups under both normal and disappearing text conditions. Error bars show standard error for each word length in each text presentation condition, for each participant group.

The 3-way interaction also reached significance (F (2, 45) = 5.67, $p < 0.001$). Paired-samples t-tests showed that for all age groups gaze durations were significantly shorter under the disappearing text condition than under the normal text condition for both long and short words (all ts > 2, all ps ≤ 0.01). However, the effect was larger for long compared to short words, and was more pronounced in younger readers compared to older readers.

Total fixation time. In contrast to gaze durations, there was no significant main effect of text presentation condition on total fixation time (F (1, 45) = 1.62, $p > 0.2$). In addition, there were no significant interactions involving text presentation condition (all Fs < 2, all ps > 0.3). There was a significant main effect of word length (F (1, 45) = 21.54, $p < 0.001$), indicating that long words elicited longer total fixation times than short words. The effect was qualified by a significant interaction with group (F (2, 45) = 6.76, $p < 0.01$), indicating a somewhat greater length effect for the younger children, though the length effect was significant for all groups (all ts > 3, all ps < 0.01) (Panel B of Figure 1).

Comparison of gaze duration and total fixation time. It is quite notable that - especially for long words - the effect of text presentation condition is highly significant in gaze duration, whereas it is reduced to the point of being non-significant in total fixation times. This later equality between conditions suggest that – in line with the global analyses – none of the age groups were disrupted by the disappearing text presentation format in reading short or long words in terms of global reading speed. However, the discrepancy between the gaze duration results and the total fixation time results indicates that all age groups adopted a different strategy in reading words, especially the longer ones, under the disappearing text condition.

More precisely, the results indicated that all age groups leave the target word quickly under the disappearing text condition (hence the gaze duration effect in favour of that condition for all groups), but return to the target words more often (hence the lack of an effect in total fixation time). It thus seems that all age groups required a second visual sample. However, when looking at the change in effect size from gaze duration to total fixation time, there does not seem to be a consistent pattern across age groups. We examined the difference between the normal and disappearing text conditions for each participant group and for each word length (Table 4).

		Effect size	
		Long target word	Short target word
Gaze duration	Adults	32 ms	23 ms
	10- to 11-years	93 ms	52 ms
	8- to 9-years	206 ms	70 ms
Total fixation time	Adults	17 ms	44 ms
	10- to 11-years	64 ms	31 ms
	8- to 9-years	-14 ms	17 ms

Table 4. The difference in reading times between text presentation conditions for the three participant groups on long and short words (time in normal text condition subtracted from time in disappearing text condition).

As can be seen in Table 4, the difference between normal and disappearing text conditions is similar in both gaze duration and total fixation time (effect sizes within 30 ms of each other) for the older children and adults on both long and short words.

In contrast, for younger children, the difference between normal and disappearing text conditions is considerably greater in total fixation times than in gaze durations, particularly for long words (effect size 220 ms greater in total fixation time than in gaze duration for long words and 53 ms greater for short words).

This suggests that a change of reading strategy when confronted with the disappearing text paradigm may not be the case for all ages, although it seems clear that the youngest children require a second visual sample of the target words, especially when they are long. That is, for the 8- to 9-year olds at least, it seems that the disappearing text paradigm forces them to leave the word sooner than they would like to, creating a need for a second visual sample (obtained through a regression back to the target word, see analyses below).

If the need for a second visual sample were the same for younger readers in the case of short as well as long words, or for older children and adults on either short or long words, one would expect this need to be reflected in the regression rate back to the target word and/or the time that is required to re-read the word, with higher regression rates and longer re-reading times for the disappearing text condition compared to the normal text condition. However, the differences in presentation condition effect size between gaze duration and total fixation time indicate that this pattern of results might only be expected for younger children reading long words. If this were the case, it would suggest that when younger children read long words they required a second visual sample. However, in the case of shorter words, and for older age groups reading both long and short words, a second visual sample was not necessarily required. We next present an analysis of refixation probability which is expected to be affected by text presentation format across age groups (cf. Blythe *et al.*, 2009. This is followed by analyses of the probability of making a regression onto the target word as well as those on re-inspection time in order to address the questions raised above.

Refixation probability. Both adults ($Z = 6.82$, $SE = 0.27$, $p < 0.001$) and older children ($Z = 3.94$, $SE = 0.21$, $p < 0.001$) made significantly fewer refixations than the younger children. More refixations were made in the normal condition compared to the disappearing text condition ($Z = 6.72$, $SE = 0.25$, $p < 0.001$), and more refixations were made on long target words than on short target words ($Z = 4.43$, $SE = 0.20$, $p < 0.001$). As in the global analysis, there was an interaction between participant group and text presentation condition – the difference between normal and disappearing text conditions was significantly smaller in adults than in children ($Z = 2.93$, $SE = 0.45$, $p < 0.01$), but was equal for older and younger children ($Z = 0.02$, $SE = 0.43$, $p = 0.99$). The interactions between word length and participant group, as well as the three-way interaction, were not significant (Zs < 1, ps > 0.5). There was a marginally reliable interaction between word length and text presentation condition ($Z = 1.72$, $SE = 0.35$, $p = 0.09$); while long target words received more refixations than short target words, this effect was greater when reading normal text compared to disappearing text (a 15%-22% difference under normal text conditions, but only about a 1%-3% difference under disappearing text conditions, see Panel C of Figure 1).

Number of regressions onto the target word. There was a close to significant main effect of text presentation (F_1 (1, 45) = 3.32, $p = 0.08$), reflecting more regressions back to target words presented as disappearing text than as normal text. However, the three-way interaction between text presentation, group and word length was significant as well (F_1 (2, 45) = 3.92, $p = 0.03$). Strikingly, as can be seen in Panel D of Figure 1, the younger children made more regressions back to the long target words in the disappearing text condition than in the normal text condition (*t*s > 3, *p*s < 0.01), while there was no significant difference between text presentation conditions for the short words (t_1 (15) = 0.53, p = 0.60). None of the t-tests comparing text presentation condition for long and short words reached significance for the older children or the adults(all *t*s < 2, all *p*s > 0.2).

Re-inspection time. Here, we subtracted gaze duration from total fixation time in order to examine the durations of regression fixations on the target words. Only trials in which a regression was made onto the target word were included in this analysis and, for this reason, there were some missing cells of data (12% of cells). These cells were filled with a global mean taken across groups and conditions. This is a conservative technique (compared to using a within-group or within-condition mean) and will have attenuated, rather than strengthened, any effects found.

There was a significant effect of group, as might be expected, showing longer re-inspection times for younger compared to older readers (F (2, 45) = 16.70, $p < 0.001$). There was also a significant effect of word length, showing longer re-inspection times on long compared to short words (F (1, 45) = 21.01, $p < 0.001$). Importantly, the interactions between group and text presentation, word length and text presentation, as well as the three-way interaction all reached significance (all *F*s > 3, all *p*s < 0.05). Younger children had longer re-inspection times in the disappearing text condition compared to the normal condition (t (15) = 2.24, p = 0.04), but there was no difference for adults or older children (both *t*s < 2, both *p*s > 0.1). Re-inspection times on long words were marginally longer in the disappearing text condition than the normal condition (t (47) = 1.93, p = 0.06) but not for short words (t (47) = 0.51, p = 0.61).

With respect to the three-way interaction, younger children had longer re-inspection times on long words in the disappearing text condition compared to the normal condition (t (15) = 2.52, p = 0.02), but there was no difference between text presentation conditions for the short words (t (15) = 0.43, p = 0.67). All other comparisons failed to reach significance (all ts < 2, all ps > 0.05). These data mirror the pattern observed in the analysis of regression probability. For the younger children, long target words received more regressions, with longer re-inspection times, when presented as disappearing text compared to normal text, but this difference did not occur for the short target words. Furthermore, this difference between reading behaviour on long words presented as normal or disappearing text did not occur for the older children or the adults.

4. Discussion

In this experiment, we compared the eye movements of adults, 10- to 11-year-old children, and 8- to 9-year-old children as they read both normal and disappearing text. A word length manipulation was included in the sentences to examine the role that refixations play in normal reading; specifically, to examine whether younger readers require a second visual sample of 8-letter words. Given our specific interest in long words, this study was conducted in Finnish - a language where long words are relatively common, even for young children.

In English, the disappearing text manipulation is minimally disruptive to readers from the age of 7-years, simply inducing a strategic change in eye movement behaviour to compensate for the fact that refixations did not provide any new information once the words had disappeared (Blythe *et al.*, 2009). However, the Blythe *et al.* study used 6-letter target words and, therefore, the conclusion that a second fixation on a word was not strictly necessary for unhindered lexical processing could well be inapplicable to longer words. Here we used the disappearing text paradigm to examine whether readers of different ages did require a second visual sample on 8-letter words compared to 4-letter words.

For fixation durations and refixation probability, the global analyses replicated the pattern reported from studies with English participants where, when reading disappearing text, participants had longer fixation durations, but made fewer refixations (Blythe *et al.*, 2009; Liversedge *et al.*, 2004; Rayner *et al.*, 2003, 2006). However, in addition to these effects, the Finnish readers from the present study also showed differences between the two text presentation conditions in other measures of eye movement behaviour. Participants also made more fixations overall and more regressions for the disappearing text compared to normal text and skipped more words. These global data certainly indicate that readers in this experiment experienced some disruption to reading when the sentences were presented as disappearing text, as there were increased numbers of fixations and regressions. Furthermore, these effects were more pronounced in younger compared to older readers. We associate this pattern of eye movement behaviour with the reader's visual sampling strategy as a consequence of the disappearing text manipulation.

The results of the local analyses (from 4 vs. 8 letter target words embedded in the experimental sentences) showed that, for younger children, eye movement behaviour on long words was very different in the disappearing text condition compared to when these words were presented normally, while there was relatively little difference between text conditions for the short words. For the 8 letter words, fewer refixations were made and gaze durations were shorter in the disappearing text condition compared to the normal condition. In contrast, more regressions were made back to these long words and their re-inspection times were longer when reading disappearing text compared to normal text. Overall, there was no difference in total fixation times. The strategy, therefore, was to initially leave the long words more quickly but then to return to them more frequently later on. This was highly effective, in that a regression provided a second visual sample of the word in a manner which had no overall cost to reading times.

Importantly, this differential pattern of visual sampling of long words when presented as disappearing text compared to normal text was present for the youngest readers (8- to 9-years old), while it was reduced for the 10- to 11-year olds, and no such difference was observed in the eye movement behaviour of adults. Thus, these data are entirely consistent with those reported by Joseph *et al.*, who found that children are more likely than adults to make refixations on long words. On the basis of those data we can form a strong argument that younger children refixate long words in order to obtain a second visual sample, and that if the opportunity to do so is denied then they adopt a different (but highly effective) strategy of regressing back to the word in order to obtain the second visual sample. The data from the present study indicate that the need for a second visual sample on eight letter words is close to adult levels by the age of 10-years. This is consistent with the data reported by Häikiö *et al.*, showing that the letter identity span extends at least seven characters to the right of fixation from the age of 10 years (Häikiö *et al.*, 2009). Thus, we interpret these data as reflecting a need in 8- to 9-year old children for a second visual sample on words of 8 letters or more, and this due to their relatively small perceptual span compared to adults.

There is, however, an alternative explanation. It may be the case that younger children's refixations on long words are language-specific, and are driven by the linguistic processing that underlies lexical identification. As mentioned in the Introduction, Finnish has a productive morphology and so long, poly-morphemic words are relatively common. On this basis, Finnish children may have learned through experience to make refixations on longer words because they are likely to be poly-morphemic, and thus they adopt a strategy that allows them additional time for morphological and/ or lexical processing. Indeed, these two explanations – perceptual span limitations leading to the need for additional visual samples, or a strategy that allows additional linguistic processing time for long words that are likely to be poly-morphemic – are not mutually exclusive, and it may be a combination of both anticipated linguistic processing demands and perceptual span limitations that lead Finnish children to make multiple fixations on long words.

The question arises, why do younger readers frequently refixate words of six letters or less (15-25% of 6 letter words), given that previous disappearing text studies have shown children do not require a second visual sample on these words (Blythe *et al.*, 2009)? One possibility is that linguistic characteristics of text such as word frequency influence refixation probability, as has been shown to be the case with skilled adult readers (e.g. Rayner, Sereno, & Raney, 1996). In comparison to work with skilled adult readers, only a few studies have examined the effect of word frequency on children's eye movements during reading and none of them reported refixation probability as a function of word frequency (Blythe *et al.*, 2009; Huestegge, Radach, Corbic, & Huestegge, 2009; Hyönä and Olson, 1995).[2]

A second possibility is that children's refixations may simply be a "habit", stemming from an earlier stage of reading development, which persists despite advances in reading skill for the individual. In this case, perhaps younger children, for example, 5- to 6-year-olds, who are within their first year or so of formal reading instruction (in the UK), need to make multiple fixations on words in order to identify them. Such a strategy might correspond to a stage of reading development where individual constituent grapheme-phoneme conversions are made in order to identify words. In this case, beginning readers may continue to make a high proportion of refixations even though their reading skill has developed beyond the point where refixations on most words are strictly necessary. Thus, when reading disappearing text, where refixations do not provide any new information, the lack of opportunity to make refixations is not detrimental to the child.

A third possibility is that refixations are a mechanism by which the reader maintains their gaze within the current word until some stage of processing with respect to that word has been achieved. In this case, high refixation probabilities may reflect the need for long gaze durations on words for beginner readers, which most naturally occur as a consequence of multiple fixations rather than a single, extended fixation. For example, it may be the case that the reader has not yet completed processing on the fixated word to a point where they are ready to move their eyes onto a new word in the sentence and thereby receive novel linguistic information to process. Maintaining gaze on the fixated word serves to prevent new information from an upcoming word interfering with processing of the fixated word. This fits well with studies showing that when reading disappearing text, the reduction in refixations is counterbalanced by an increase in fixation durations (Blythe *et al.*, 2009; Liversedge *et al.*, 2004; Rayner *et al.*, 2003). Thus, it may be that the reader requires a certain amount of processing time on the fixated word and they achieve this by one of two mechanisms – refixations or longer initial fixation durations – depending on whether there is an attended visual object available to refixate (normal presentation) or if the attended word has disappeared.

It seems likely that there will not be a single account underlying children's refixations in reading; rather, it may be a combination of these (and possibly other) alternatives. For example, research with skilled adult readers has shown that orthographic characteristics can influence the location of refixations (White & Liversedge, 2004, 2006a, 2006b). Additional research may well show that linguistic and orthographic characteristics of text also influence refixation probability in children as well as adults (though note that the locations of children's refixations are non-systematic compared to those of adults, see Joseph *et al.*, 2009).

The final issue to consider is whether these results can be generalized to other languages. This is particularly pertinent, given that the key manipulation in this experiment was one of word length and that Finnish readers are far more familiar with long words when they read compared to English readers. This could, theoretically, lead to some language-specific characteristics of eye movement behaviour, reflecting differences in cognitive processing with respect to long words. First, it should be noted that other Germanic languages, such as German, Dutch, and Swedish, are similar to Finnish in that they all contain relatively many long words compared to English. Thus, the concern becomes, more specifically, can the results from the present study be generalized to languages such as English which do not contain so many long words? There are two main areas of research which support the argument that, despite differences in the commonality of long words, there is no reason to believe that the results of the present experiment do not generalize to English.

First, with respect to the perceptual span in children and adults, it should be noted that studies in both English and Finnish have found highly similar patterns of results despite differences between the languages (Häikiö *et al.*, 2009; Rayner, 1986). Moreover, the present experiment shows a highly similar pattern of differences between age groups to that previously observed in studies with English readers (Blythe *et al.*, 2009). Importantly, all these studies show that the development of visual encoding during fixations in reading, with respect to both spatial and temporal aspects of encoding, develops similarly in both Finnish and English readers.

Second, the data from the normally presented text condition here show a similar pattern to the data reported by Joseph *et al.*, who used a word length manipulation with adult and child readers in English (although note that the words in that study had far lower frequencies than those in the present study). The Joseph *et al.* study included children aged between 7- and 11-years. In both that study and the present study, 8-letter and 4-letter words were compared. The difference in mean gaze durations between long and short words was 16 ms for adults and 88 ms for children in English; in Finnish, the difference was 15 for adults, 50 ms for 10- to 11-year-olds, and 210 ms for 8- to 9-year-olds. The mean difference in refixation probability was 0.11 for adults and 0.24 for children in English; in Finnish, the difference was 0.15 for adults, 0.15 for 10- to 11-year-olds, and 0.21 for 8- to 9-year-olds. Thus, the data from these two studies do not give any reason to think that readers of any age process long compared to short words differentially depending on the language spoken. Rather, the data from all of the studies reported here show strong similarities between studies in both English and Finnish with respect to visual sampling and information encoding during fixations in reading; thus, it seems reasonable to conclude that the results from this study can be generalized to English.

5. References

Blythe, H.I., & Joseph, H.S.S.L. (2010). Children's eye movements during reading. In S. Liversedge, I. Gilchrist, and S. Everling (Eds.) Oxford Handbook of Eye Movements. *Invited chapter, accepted subject to minor revisions*.

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.

Brysbaert, M., Drieghe, D., & Vitu, F. (2003). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.) *Cognitive Processes in Eye Guidance.* Oxford University Press: Oxford, UK.

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.

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., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? Evidence from an eye-movement contingent display change study. *Memory & Cognition, 32,* 523-532.

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 effects and landing positions during reading in children and adults: Evidence from eye movements. *Vision Research, 49,* 2078-2086.

Juhasz, B.J. (2008). The processing of compound words in English: Effects of word length on eye movements during reading. *Language and Cognitive Processes, 23,* 1057-1088.

Kirkby, J.A., Webster, L.A.D., Blythe, H.I., & Liversedge, S.P. (2008). Binocular coordination during reading and non-reading tasks. *Psychological Bulletin, 134,* 742-763.

Laine, M., & Virtanen, P. (1999). *WordMill Lexical Search Program*. Center for Cognitive Neuroscience, University of Turku, Finland.

Liversedge, S.P., Rayner, K., White, S.J., Vergilino-Perez, D., Findlay, J.M. & Kentridge, R.W. (2004). Eye movements when reading disappearing text: Is there a gap effect in reading? *Vision Research, 44,* 1013-1024.

McConkie, G.W., Kerr, P.W., Reddix, M.D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixation locations on words. *Vision Research, 28,* 1107-1118.

Raaijmakers, J.G.W., Schrijnemakers, J.M.C., & Gremmen, F. (1999). How to Deal with "The Language-as-Fixed-Effect Fallacy": Common Misconceptions and Alternative Solutions. *Journal of Memory and Language, 41,* 416 – 426.

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 and information processing: 20 years of research. *Psychological Bulletin, 124,* 372-422.

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., Liversedge, S.P., & White, S.J. (2006). Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research, 46,* 310-323.

Rayner, K., Sereno, S.C., & Raney, G.E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1188-1200.

Rayner, K., Yang, J., Castelhano, M.S., & Liversedge, S.P. (2010). Eye movements of older and younger readers when reading disappearing text. *Psychology and Aging, under revision.*

Shen, D., Liversedge, S.P., Tian, J., Bai, X., Yan, G., & Rayner, K. (2010). Second language learning: Reading spaced and unspaced text. *In prep.*

White, S.J., & Liversedge, S.P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16,* 52-78.

White, S.J., & Liversedge, S.P. (2006a). Foveal processing difficulty does not modulate non-foveal orthographic influences on fixation positions. *Vision Research, 46,* 426-437.

White, S.J., & Liversedge, S.P. (2006b). Linguistic and non-linguistic influences on the eyes' landing positions during reading. *The Quarterly Journal of Experimental Psychology, 59,* 760-782.

Footnotes

[1] A full set of experimental materials can be obtained by contacting the first author.

[2] Given the outstanding question of which factors underlie children's refixations during reading, we decided to examine the effect of word frequency on children's refixation probability using the Blythe *et al.* data set. Across two experiments, high and low frequency target words were embedded in sentences that were presented either normally or as disappearing text. In Experiment 1, adults and children aged 7- to 11-years were compared, with a 60 ms delay in the disappearing text condition. In Experiment 2, adults, 10- to 11-year olds and 7- to 9-year-olds were compared, using three different disappearing text delays. Here, given the more general nature of the question of which factors underlie children's refixations in reading, for simplicity's sake we only present the data from the normally presented text conditions in these two experiments.

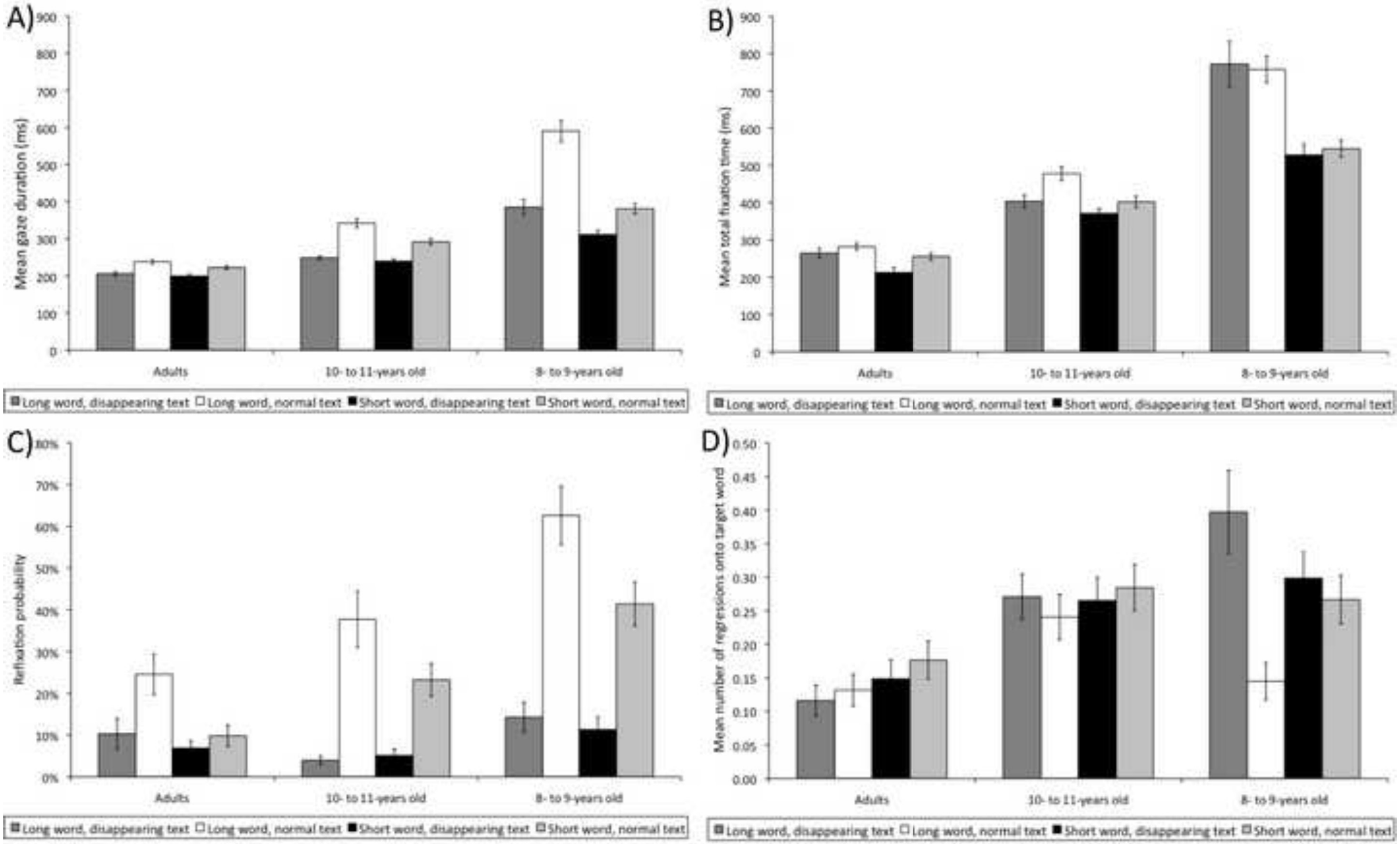
In Experiment 1, there was no hint of an effect of word frequency on refixation probability for either adults or children. Adults refixated 12% of high frequency words and 13% of low frequency words (t_1 (11) = 0.17, p = 0.87; t_2 (39) = 0.11, p = 0.91), while children refixated 35% of high frequency words and 33% of low frequency words (t_1 (11) = 0.11, p = 0.92; t_2 (39) = 0.65, p = 0.52). In Experiment 2, there was a numerical trend in the data for adults to refixate low frequency words more often than high frequency words (5% and 13%, respectively), but this difference was not reliable (t_1 (15) = 1.86, p = 0.08; t_2 (39) = 1.64, p = 0.11). For both groups of children there was very little difference in refixation probability between high and low frequency words although the numerical differences were in the same direction as the trend in the adult data. Older children refixated 21% of high frequency words and 25% of low frequency words (t_1 (15) = 0.68, p = 0.51; t_2 (39) = 0.57, p = 0.57), while younger children refixated 24% of high frequency words and 28% of low frequency words (t_1 (15) = 0.57, p = 0.58; t_2 (39) = 0.43, p = 0.67).

Thus, in the Blythe *et al.* data set there is no strong evidence for an effect of word frequency on refixation probability. We would expect to see such an effect in the adult data, on the basis of previous studies (such as Rayner *et al.*, 1996). Further research is clearly needed in order to directly investigate the role of cognitive processing difficulty on refixation probability in beginning compared to skilled readers.

Acknowledgements

This research was supported by Grants F/00 128/AB and F/00 128/AG from the Leverhulme Trust, and Grant G0601757 from the MRC to S. Liversedge. H. Blythe was supported by an ESRC Postdoctoral Fellowship. R. Bertram was supported by a grant of the Academy of Finland (Grant #118404). We would like to thank SR Research for programming assistance, and the staff and children from the primary school who supported this experiment with such enthusiasm. We would also like to thank Erik Reichle and an anonymous reviewer for their helpful comments.

Figure



A)
Mean gaze duration (ms)
0
100
200
300
400
500
600
700
800
900
Adults
10- to 11-years old
8- to 9-years old
Long word, disappearing text
Long word, normal text
Short word, disappearing text
Short word, normal text
B)
Mean total fixation time (ms)
C)
Refixation probability
0%
10%
20%
30%
40%
50%
60%
70%
80%
D)
Mean number of regressions onto target word
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50