**Is a frequency effect observed in eye movements during text reading? A comparison
between developing and expert readers**

Rachid Khélifi, Laurent Sparrow, Séverine Casalis

Laboratoire SCALab UMR CNRS 9193,

Lille 3 University, France

Corresponding author:

Rachid Khélifi

Laboratoire SCALab UMR CNRS 9193,

Université Lille 3, BP 60149, 59653,

Villeneuve d'Ascq Cedex, FRANCE

E-Mail : rachid.khelifi@univ-lille3.fr

Acknowledgements

We thank the anonymous reviewers for insightful comments and suggestions.

We would like also to thank Dr Solène Kalénine for her support in writing this paper.

Abstract

In the present study, the frequency effect was examined specifically in text reading in developing and expert readers. Third graders, fifth graders and expert readers read the same short text containing high- and low-frequency words while their eye movements were recorded. The frequency effect was observed in temporal and spatial measures of eye movement in all participants. However, we found a greater frequency effect in developing readers than in expert readers, and also in third graders than in fifth graders. Furthermore, the frequency effect emerged at an early stage, as soon as the first fixation duration in third graders and gaze duration in fifth graders. In contrast, the frequency effect was observed only for total fixation in expert readers. We discuss these results within the framework of the computational model E-Z Reader to explain why the magnitude of the frequency effect could be different across reading levels.

1. Introduction

Over the last 40 years, a large body of research has examined the cognitive processes involved in skilled reading using eye-tracking technology (for a review, see Rayner, 1998; 2009). Eye movement recording has become a major instrument in revealing the moment-to-moment activity of the mind during reading. With the accumulation of knowledge, several computational models have been proposed to account for different mechanisms – perceptual, attentional and oculomotor – underlying eye behavior in reading (Engbert, Nuthmann, Richter, & Kliegl, 2005; Legge, Hooven, Klitz, Mansfield, & Tjan 2002; Reichle, Pollatsek, Fisher, & Rayner, 1998). Despite major advances in the area of expert reading, interest in a developmental approach to eye movement is very recent (Miller & O'Donnell, 2013; Schroeder, Hyönä, & Liversedge, 2015). The basic differences in eye movement behavior between developing and expert readers are now well known: developing readers have a longer fixation duration, carry out more fixations and refixations, and make shorter saccades than expert readers (for a review, see Blythe & Joseph, 2011; Radach, Schmitten, Glover, & Huestegge, 2009). However, there is still a lack of knowledge about the developmental trajectory by which children become skilled readers.

In expert readers, many eye movement studies have focused on the frequency effect (Henderson & Ferreira, 1990, 1993; Inhoff, 1984; Rayner & Duffy, 1986; Raney & Rayner, 1995; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989; Schilling, Rayner, & Chumbley, 1998). It has been shown that high-frequency words are fixated for a shorter length of time, more often skipped – that is, not fixated – and are less refixated than low-frequency words (Inhoff, 1984; Rayner & Duffy, 1986; Rayner & Liversedge, 2011; Schilling et al., 1998; Vitu, 2005). The frequency effect in eye movement has been also reported in developing readers (Huestegge, Radach, Corbic, & Huestegge, 2009; Hyönä & Olson, 1995; Joseph, Nation, & Liversedge, 2013; Rau, Moeller, & Landerl, 2014; Tiffin-Richards & Schroeder,

2015; Vorstius, Radach, & Lonigan, 2014). Moreover, the effect was greater in developing readers than in expert readers (Joseph et al., 2013; Tiffin-Richards & Schroeder, 2015). It is difficult to compare developing and expert readers directly since, in most studies, expert readers are university students. However, not all primary school pupils will attend university so comparisons are hazardous. Nevertheless, both groups are exposed daily to written texts. It is more common to use text rather than independent sentences in expert readers than in beginning readers (in expert readers: Chamberland, Saint-Aubin, & Légère, 2013; Cop, Keuleers, Drieghe, & Duyck, 2015; Miellet, Sparrow, & Sereno, 2007; Raney & Rayner, 1995; Rayner & Raney, 1996; Wang, Pomplun, Chen, Ko, & Rayner, 2010; Whitford & Titone, 2012; in developing readers: De Leeuw, Segers, & Verhoeven, 2016; Hyönä & Olson, 1995; Valle, Binder, Walsh, Nemier, & Bangs, 2013). However, text reading and single-sentence reading are different tasks. Indeed, specific differences between them have already been reported in eye movement studies. For example, Huestegge and Bocianski (2010) showed that in comparison to single sentence reading, eye movement on a current sentence in text reading can be modified by a sentence that has been previously read. Furthermore, frequency may have a different impact according to whether words are presented in isolation, in single sentences or in text. For example, Wallot, Hollis and van Rooij (2013) found that the influence of lexical variables such as word frequency, word length and word co-occurrences on reading times explained roughly 2% of the variance in text reading, while it explained from 15% to 35% of the variance in reading single words or single sentences in expert readers (Baayen, 2010; Cordier, Croizet, & Rigalleau, 2012).

Consistent with the previous results, Radach, Huestegge and Reilly (2008) reported that the size of the frequency effect was reduced when the same target words were read in a paragraph rather than in single sentence. Moreover, the reduction was greater in gaze duration and total reading time measures than in first fixation duration or in single fixation

measures. The authors suggested that a reduction in frequency effect was due to the higher order post-lexical processing required by text reading, e.g. integrating the meaning of a word within a mental representation. Text reading involves a more complex process of constructing meaning than single sentence reading. Indeed, when reading a text, the reader has to pool different types of information from the processing of several sentences, the aim being to construct a consistent representation of the text that is read.

Thus, the frequency effect observed in text reading may not be similar in magnitude to that reported in single word or single sentence reading, given the additional processing required by text reading. In other words, the frequency effect may be less easy to evidence in text reading than in sentence reading. For this reason, frequency effects should be examined in children specifically during text reading and should be compared with those observed in expert readers. These are the main goals of our study.

Concerning the different measures used in eye movement studies, temporal measures should be distinguished from spatial measures (Radach & Kennedy, 2004). Temporal measures usually comprise the first fixation duration (duration of the first fixation within a word), the gaze duration (summed duration of all fixations made on a word before leaving it), the total reading time (summed duration of all fixations made on a word) and the duration of single fixations (duration of the fixation on a word that is fixated only once). It is also generally agreed that first fixation, gaze duration and single fixation reflect lexical access processes, while the total reading time corresponds to both initial and later processing activities (Rayner, Juhasz, & Pollatsek, 2005). Spatial measures usually express the probability of fixation (i.e. the probability that a word is fixated at least once, the inverse measure being the skipping rate, i.e. the probability that a word is not fixated), the probability of refixation (i.e. the probability that a word is fixated more than once) and the saccade length (i.e. the distance in character space between the mean position of two successive fixations).

Studies that have explored the frequency effect in developing readers have mainly used two methods: reading aloud (Huestegge et al., 2009; Hyönä & Olson, 1995) versus silent reading (Blythe et al., 2009; Huestegge et al., 2009; Joseph et al., 2013; Tiffin-Richards & Schroeder, 2015). Moreover, some have used texts including words from age-appropriate databases (De Leeuw et al., 2016; Hyönä & Olson, 1995; Joseph et al., 2013; Tiffin-Richards & Schroeder, 2015; Valle et al., 2013; Vorstius et al., 2014) or adult corpus data to index word frequency (Blythe et al., 2006; Blythe et al., 2009; Huestegge et al., 2009). In addition, while most studies used the same text that was generally appropriate to the children's age, some used different texts so each age group had to read a text that was age-appropriate. For example, Hyönä and Olson (1995) asked young readers aged 10.5 years on average to read aloud two 10-line texts containing the target words classified as high-, medium-, or low-frequency. They demonstrated a frequency effect in first fixation duration as well as in gaze duration: fixation durations increased as word frequency decreased. Moreover, low-frequency words received more fixations than high-frequency words. De Leeuw et al. (2016) reported that in fourth graders, high-frequency words were skipped more often and had shorter gaze durations than low-frequency words in text reading. However, rather than being experimentally manipulated, word frequency in the two previous studies was assessed post hoc, thus introducing a possible confusion with other linguistic factors such as word length (Rayner & Duffy, 1986). More recently, Huestegge et al. (2009) addressed this methodological bias in a study conducted in second and fourth graders. Participants read aloud sentences containing child-appropriate statements with embedded target words that varied orthogonally in length (short vs. long) and frequency (high vs. low). Overall, the authors observed a classic frequency effect that was greater in second than in fourth graders. However, in contrast to Hyönä and Olson's result (1995), a frequency effect appeared later and was observed only in gaze duration and total reading time.

In the studies by both Huestegge et al. (2009) and Hyönä & Olson (1995), the participants were assessed in the oral reading paradigm. However, reading modality (e.g. aloud vs. silent) may have a significant influence on the frequency effect. For example, Vorstius et al. (2014) showed that the frequency effect in developing readers tended to be less pronounced in oral reading than in silent reading. In other words, oral reading is probably not the most relevant paradigm to highlight a frequency effect in developing readers. Blythe et al. (2009) examined the frequency effect in both developing and expert readers who were required to read sentences silently, each of them containing a high- or low-frequency target word. A frequency effect was observed in both types of readers in first fixation duration, gaze duration and single fixation duration. Nevertheless, this effect was not greater in developing readers than in expert readers. Both Huestegge et al. (2009) and Blythe et al. (2009) classified target words as high- or low-frequency in reference to CELEX, a lexical database of adult written language. However, it is essential to control word frequency carefully when it is examined in children. In fact, some words that are classified as high-frequency for adults are found to be low-frequency for children. As frequency count depends on the number of times a word is encountered, it is clear that the same word may have a different frequency count according to age. Thus, a rigorous examination of a frequency effect in children must use age-appropriate word frequencies for children. This issue was addressed by Tiffin-Richards and Schroeder (2015), Joseph et al. (2013) and Valle et al. (2013). The study by Tiffin-Richards and Schroeder (2015) was conducted in second-grade developing readers and in expert readers. The participants read silently a set of sentences in which a high- or low-frequency word was embedded. The authors found a frequency effect in first fixation duration, gaze duration and total reading time in developing readers. In contrast, expert readers showed a frequency effect only in first fixation duration. In the same vein, Joseph et al. (2013) asked third graders and expert readers to read a set of sentences, each containing a target word of high or low

frequency. Moreover, the targets were controlled according to the adult norm of word frequency. In developing readers, a frequency effect was observed in gaze duration and total reading time, while in expert readers a frequency effect was found only in total reading time. The results reported in the two previous studies are quite different, even though both of them used an isolated sentence reading task and a child corpus to index frequency. In developing readers, a frequency effect emerged earlier (first fixation duration) in the study by Tiffin-Richards and Schroeder (2015) than in that by Joseph et al. (2013). Furthermore, both studies reported a greater frequency effect in developing readers than in expert readers.

To our knowledge, only one study has examined the frequency effect on developing readers in text reading (Valle et al., 2013). In this study, second graders had to read a text composed of 157 words including high- and low-frequency words. Results evidenced a frequency effect in gaze duration and in total reading time but not in the first fixation duration.

In all, our understanding of the frequency effect in various reading situations in children (word, sentence and text) is far from complete. We specifically need information on how the frequency effect of text reading evolves during reading acquisition.

2. Aims of the present study

Most of the abovementioned studies conducted in developing readers reported a frequency effect in sentence reading. However, the time at which a frequency effect began to emerge differed across those studies. In fact, even though child corpus data were used to index word frequency, a frequency effect was systematically reported in gaze duration (Joseph et al., 2013; Tiffin-Richards & Schroeder, 2015) but not in first fixation duration (Joseph et al., 2013) in sentence reading. To our knowledge, only a few eye movement studies have used text reading to examine the frequency effect in developing readers (De Leeuw et al., 2016; Hyönä & Olson, 1995; Valle et al., 2013). Among these, only one involved a silent reading task with an experimental manipulation of frequency (Valle et al., 2013) and it included only

second graders. The absence of more advanced readers (older children or even expert readers) does not provide a developmental view of the evolution of the frequency effect in text reading. Therefore, the aim of the current research was to examine the frequency effect in silent text reading in two groups of developing readers, one of third graders and one of fifth graders, and also to compare their results to those collected in expert readers. To assess more precisely the impact of word frequency, a child frequency count was used. Lastly, the research presented in the introduction was based largely on temporal measures. In our study, we also wanted to examine the frequency effect in terms of both temporal and spatial measures. Considering the results previously reported, we had several hypotheses. First, we expected that a frequency effect would be observed in both developing and expert readers, the effect being evidenced in temporal measures with longer fixation durations for low- than in high-frequency words. As reported by Vorstius et al. (2014), we also expected that it would be evidenced in spatial measures with higher fixation and refixation probabilities in low- than in high-frequency words. Second, as noted in previous studies (Joseph et al., 2013; Tiffin-Richards & Schroeder; 2015), we expected a larger frequency effect in developing readers than in expert readers.

Indeed, even if text reading involves more high-level processing than the reading of sentences or isolated words, lexical processing is still very important in children because it is not as automatic as in expert readers. As reported by a recent study (Joseph & Livversedge, 2013), high-level processing tends to be delayed in children in comparison to expert readers, a result which is consistent with the idea that lexical processing is the core issue in low proficient readers (Perfetti, 2007). Thus, we predicted that the less skilled the reader is, the greater the frequency effect would be. In other words, the frequency effect should be greater in third graders than in fifth graders because third graders are supposed to be less experienced and to know fewer words than fifth graders.

3. Materials and methods

3.1. Participants

We recruited 29 third graders (age: $M = 8.4$ years, 14 girls), 30 fifth graders (age: $M = 10.8$ years, 15 girls) and 25 adults (age: $M = 23.6$ years, 12 girls). Children were recruited from five primary schools in the area of Lille, northern France. Adults were students at Lille University. All participants were native French speakers with normal or corrected-to-normal vision and no known reading difficulties. Before participation in the experiment, written informed consent was obtained from each child's parent or legal guardian and written assent was obtained from each child. Written informed consent was also obtained from the expert readers. The children's reading level was assessed using the French Alouette test (Lefavrais, 1967), a standardized task that consists of reading a 265-word text as quickly and accurately as possible. The number of words correctly read and the reading time are transformed into a reading-age equivalent. None of the children had a reading lag above 12 months, with 18 months being the maximal delay to be considered a normal reader. Their nonverbal reasoning scores were also tested using the Raven's Progressive Matrices (Raven, 1947) and were within the normal range. Chronological age, reading age and nonverbal reasoning scores are reported in Table 1.

3.2. Material

For the purpose of this research, a text containing 127 words was used ("L'oiseau et le cheval", written by Khélifi Abdel-Kader, see Appendix A). This text was produced from the selection of 20 words from the Manulex-Infra database (Peereman, Lété, & Sprenger-Charolles, 2007). This database lists the word-frequency values for 1.9 million words extracted from 54 French elementary school books between the first and fifth grades. While a corpus of at least 20 million words also including subtitles to estimate word frequency is recommended (Brysbaert, Buchmeier, Conrad et al., 2011), we used the Manulex-infra based

on textbooks because this is the only database available in French for children. Only words corresponding specifically to the third-to-fifth grade readers level were selected. Ten words were high-frequency, with a printed frequency of between 90 and 1277 occurrences per million ($M = 348$), and 10 words were low-frequency, with a printed frequency of between 13 and 50 occurrences per million ($M = 28$). High-frequency and low-frequency words were matched on word length (which ranged from 4 to 9 letters). We also reported word age-of-acquisition (AOA) by using Gilhooly and Logie's (1980) rating methods. Table 2 lists different features of the high- and low-frequency target words. Other words used in the text were also selected from the same abovementioned database. The text described a short story that is easy to understand for developing readers from grade 3. Target words were distributed homogeneously throughout the text and none of them appeared at the beginning or end of a line. Target words were also controlled according to the Lexique database (New, Pallier, & Ferrand, 2004), an adult frequency count. Mean frequency in written language was 271 per million for high-frequency words (range: 50-1289) and 21 per million for low-frequency words (range: 1-39). The predictability (probability with which a word can be guessed from the prior context) of each word of the text was assessed. To this end, 20 third graders, 20 fifth graders and 24 expert readers who did not take part in the main study were recruited in the same context as previously (native French speakers with no reading difficulties). At the beginning of the assessment, no word was displayed and participants were required to guess the first word of the text. After giving their response, the first word was displayed and the second word had to be guessed from the preceding context. This procedure was repeated until the last word of the text. Predictability was 0.10, 0.15 and 0.27 in third graders, fifth graders and expert readers, respectively. While predictability rates did not differ between third and fifth graders, it was significantly higher in expert readers than in both groups of developing readers, all $ps < .001$. Furthermore, low-frequency words were more predictable than high-

frequency words in expert readers, but not in third or in fifth graders. For target words, predictability is presented in Table 2 as a function of frequency.

3.3. Procedure

Participants were tested in a quiet room at school for developing readers and in a university laboratory for expert readers. Both places were protected from the sunlight and illuminated by an indirect 13-W compact fluorescent lamp to ensure similar lighting conditions for all participants. Experiments took place between January and March in the relevant school year. Participants were required to read a text silently on a screen. The experiments were conducted with the E-prime 2.0 software (Psychology Software Tools, Sharpsburg, PA, USA) running on a Windows computer. An ASL EYE-TRAC 6000 system was used to record participants' eye movements as they read the text. Viewing was binocular but the fixation location was monitored only from the right eye. The subject's eyes were illuminated by an invisible infra-red light. Pupil diameter and corneal reflections were recorded by a video-camera. A line of gaze was extracted at a sampling rate of 60 Hz¹. A chin-rest was used to minimize head movements and maintain a 65 cm viewing distance from the screen. One character subtended approximately 0.34 ° of visual angle.

Prior to each session, the eye-tracker was calibrated using a standard nine-point calibration displayed on the computer screen. The nine points were arranged in three rows and three columns. Participants were required to fixate successively each of the nine points. If necessary, the calibration procedure was repeated until the pupil and corneal reflection could be detected by the system for all nine points. After the calibration stage, text reading took place. The text was divided into four paragraphs consisting of three to six sentences. Each paragraph was displayed successively on the screen. Before this, participants were required to

¹ As reported by Raney, Campbell, and Bovee (2014), a 60 Hz sampling rate is sufficient for recording fixation duration.

fixate a white dot appearing at the upper-left corner of the screen. Once the dot was fixated, the experimenter started the display of a paragraph. After reading the paragraph, participants were required to fixate a red dot appearing at the lower-right corner of the screen. They were instructed to read as naturally as possible in order to understand the story. They were also informed that after the text reading, they would have to answer some comprehension questions orally. A complete testing session lasted approximately 15-20 minutes. After the text had been read, its comprehension was briefly assessed by five questions to which the participant responded orally. Answers were recorded by the experimenter. Incorrect answers were scored 0, partially correct answers 1 point and correct answers 2 points. The maximum comprehension score was thus 10 points. Results are reported in Table 1. All participants had a score between 7 and 10. Third graders ($M = 8.27$) and fifth graders ($M = 8.50$) did not differ on comprehension score ($t < 1, p = .46$). At the end of the session, reading level was evaluated with the Alouette test and nonverbal reasoning was assessed in beginning readers.

3.4 Data analysis

Various measures were extracted using an ad-hoc program in MATLAB. For temporal measures, first fixation duration, gaze duration, total reading time and single fixation duration were reported. For spatial measures, both fixation and refixation probabilities were considered. We also reported the percentage of single fixation. Prior to analysis, all fixations with a blink or with a duration shorter than 80 ms were discarded. In addition, first fixations greater than 1200 ms, gaze durations greater than 2400 ms, or total viewing times longer than 4800 ms were excluded from the analyses (8.5% of the initial data).

Values of temporal measures did not fit a normal distribution. However, log-transformation used to normalize the data may affect the interpretation of results (Lo & Andrews, 2015). Indeed, statistical analysis is conducted on log-transformed data while interpretation is done on back-transformed data, i.e. on the original untransformed data. There is no guarantee that

the effect observed on log-transformed data would be the same on the original untransformed data. To avoid this difficulty, we decided to use general linearized mixed models (GLMMs), which, unlike linear mixed models, are appropriate when data does not meet a normal assumption. Given that the distributions of fixation duration were skewed to the right, we decided to use the gamma family (which is appropriate for skewed data) and the identity link (because we consider that fixation duration is a direct measure of word identification). To examine spatial binary measures, we also used GLMMs with the binomial family and logit link function. Post hoc comparisons were conducted with the `emmeans()` function.

4. Results

Statistical analyses were conducted using the `lme4` package (version 0.999999-4; Bates, Maechler, & Dai, 2013) within the R Environment for Statistical Computing (R Development Core Team, 2013). Both group and frequency were entered as fixed effects while participants and items were entered as random effects. To examine the significance of individual and interacting factors, the function `join-test()` was used. When the interaction between frequency and group was significant, further analysis were carried out in each group of participants. We used the `summary()` function to examine whether the frequency effect was significant. In this case, frequency was entered as fixed effect while participants and items were entered as random effects. Table 3 shows the mean fixation duration for temporal measures and Table 4 the mean fixation probability for spatial measures.

4.1 Analysis of temporal measures in beginning and expert readers

First fixation duration

There was a significant effect of group, $F = 6.83, p < .001$, and a significant effect of frequency, $F = 14.91, p < .001$. Interaction between group and frequency was also significant, $F = 6.12, p < .01$. Effect of frequency was significant in third graders, ($b = 0.093, SE = 0.020, t = 4.64, p < 0.001$, Cohen's $d = 0.47$), did not reach significance in fifth graders, ($b = 0.032,$

$SE = 0.018$, $t = 1.74$, $p = 0.082$, Cohen's $d = 0.16$) and was clearly not significant in expert readers, $t < 1$, $p = .89$, Cohen's $d = 0.03$. Thus, first fixation duration was longer in low- than in high-frequency words in third graders, but not in fifth graders or in expert readers.

Percentage of single fixation and single fixation duration

The percentage of single fixation (the ratio between the total number of single fixations and the total number of multiple fixations) was 44%, 57% and 68% in third graders, fifth graders and expert readers, respectively. Effect of group was significant, $F = 26.67$, $p < .001$ as was effect of frequency, $F = 14.10$, $p < .001$. Percentage of single fixation was significantly lower in third graders ($M = 0.45$) than in fifth graders ($M = 0.58$) and expert readers ($M = 0.69$), all $ps < .001$. It was also lower in fifth graders than in expert readers, $p = .031$. Percentage of single fixation was higher in high- ($M = 0.62$) than in low-frequency words ($M = 0.52$).

Interaction between group and frequency was not significant, $F < 1$.

For single fixation duration, group effects, $F = 11.03$, $p < .001$, and frequency effects, $F = 4.99$, $p < .05$, were significant. Interaction between group and frequency was not significant, $F = 1.82$, $p = .160$. This nonsignificant interaction suggests that the frequency effect was present in all participants. However, this result should be considered cautiously because of the lower rate of single fixation in third graders than in the other groups.

Gaze duration

Effect of group was significant, $F = 14.73$, $p < .001$, as was the effect of frequency, $F = 25.27$, $p < .001$. There was a significant group $\times$ frequency effect, $F = 3.99$, $p < .05$. Gaze duration was longer in low- than in high-frequency words in third graders ($b = 0.096$, $SE = 0.115$, $t = 3.74$, $p < 0.001$, Cohen's $d = 0.42$) and in fifth graders ($b = 0.071$, $SE = 0.023$, $t = 3.10$, $p < 0.01$, Cohen's $d = 0.32$) but not in expert readers ($b = 0.026$, $SE = 0.019$, $t = 1.36$, $p = 0.18$, Cohen's $d = 0.21$).

Total fixation duration

There was a significant effect of group, $F = 24.62, p < .001$, and a significant effect of frequency, $F = 28.41, p < .001$. Interaction between group and frequency was also significant, $F = 4.30, p < .05$. Total reading time was longer in low-frequency than in high-frequency words in third graders ($b = 0.195, SE = 0.047, t = 4.11, p < 0.001$, Cohen's $d = 0.43$), in fifth graders ($b = 0.077, SE = 0.032, t = 2.39, p < 0.05$, Cohen's $d = 0.26$) and in expert readers ($b = 0.058, SE = 0.028, t = 2.08, p < 0.05$, Cohen's $d = 0.24$).

4.2 Analysis of temporal measures in developing readers only

In the previous analysis, the interaction between frequency and group in first fixation duration, gaze duration and total reading time suggested that that frequency effect was modulated according to the reading level. However, this significant interaction may have resulted mainly from an absence of frequency effect in expert readers. To examine this possibility, we conducted an additional analysis by considering only developing readers. Interaction between group and frequency was significant in first fixation, $F = 5.02, p < .05$, and total reading time, $F = 4.98, p < .05$ but not in gaze duration, $F = 1.51, p = .22$. Thus, the frequency effect was greater in third graders than in fifth graders in both first fixation duration and total reading time.

4.3 Analysis of spatial measures

Fixation probability

There was a significant effect of group on fixation probability, $F = 8.20, p < .001$ and a significant effect of frequency, $F = 4.54, p < .05$. Interaction between group and frequency was not significant, $F < 1$. Post hoc comparisons revealed that the fixation probability for both high- and low-frequency target words was higher in third ($M = 0.66$) and fifth graders ($M = 0.70$) than in expert readers ($M = 0.59$), all $ps < .05$. In contrast, no significant difference was found in fixation probability between third and fifth graders, $p = .31$.

Refixation probability

There was a significant effect of group on refixation probability, $F = 15.26, p < .001$ and a significant effect of frequency, $F = 26.87, p < .01$. Interaction between group and frequency was not significant, $F < 1$. Pairwise comparisons revealed that the refixation probability in both high- and low-frequency words was higher in third ($M = 0.18$) and fifth graders ($M = 0.14$) than in expert readers ($M = 0.09$), and higher in third graders than in fifth graders, all $ps < .05$.

5. Discussion

The purpose of the present study was to investigate the frequency effect in developing readers during text reading and to compare the results with those of expert readers. A silent reading task was used and we recorded eye movements. The frequency count of target words was derived from an age-appropriate database. To examine whether the frequency effect differed according to reading level, third graders and fifth graders participated in the experiment. Consistent with the few previous studies that have examined the question with an adequately controlled methodology (Tiffin-Richards & Schroeder, 2015; Joseph et al., 2013; Valle et al., 2013), our results evidenced a frequency effect in text reading in temporal measures: all participants fixated low-frequency words for a longer time than high-frequency words (at least in the total reading time).

An important issue of the present study was to examine the impact of age on frequency effects to track effects related to reading development. Interestingly, we found an interaction between age group and frequency on first fixation, even when the expert group was removed. This clearly indicates that the impact of frequency decreases with proficiency. Note that while we also evidenced an age effect on total fixation time, the impact on gaze duration was not found to be stronger in younger readers.

Considering the spatial measures, frequency influenced both fixation and refixation probabilities for all participants: low-frequency words were fixated more and also received

more multiple fixations than high-frequency words in both developing and expert readers. Therefore we observed developmental changes in frequency effects in temporal measures but not in spatial measures. Importantly, the frequency effect became significant from first fixation and gaze measure in developing readers, while it was only for total fixation duration in expert readers. It thus appears that the less skilled the readers are, the earlier the frequency effects emerge in fixation duration. This study is the first to consider two groups of developing readers and to evidence the impact of proficiency and reading exposure to frequency effects in text reading. Therefore, our results replicate previous findings indicating that the frequency effect survives in text reading in developing children, and extend them by showing a quasi-linear effect of proficiency, with early effects emerging for less skilled readers.

Unlike us, however, previous authors evidenced a frequency effect in expert readers that was usually observed from the first fixation duration (Miellet et al., 2007; Raney & Rayner, 1996; 1995). This issue is discussed below.

We interpret our results within the framework of the computational model E-Z Reader (Reichle, Rayner, & Pollatsek, 2003), which explains how word identification, visual processing, attention, and the oculomotor system work together to produce the patterns of eye movements observed in reading. It considers that words are identified in two steps: (i) an early stage of lexical processing (*L1*) called the familiarity check, in which word familiarity is rapidly assessed; (ii) a later stage called (*L2*), corresponding to the completion of the lexical access. The distinction between these two stages is based on the assumption that *L1* allows recovery of the word-form information (e.g. orthographic and/or phonological) while *L2* allows recovery of its meaning (Rayner et al., 2003).

How can E-Z Reader account for the frequency effect and for the fact that it is greater in the youngest readers and more generally in developing readers than in expert readers? The

frequency effect can impact word processing in two ways. First, when a word is encountered, the first stage assesses its familiarity. It takes more time to check the familiarity of low-frequency words encountered only a few times than well-known high-frequency words encountered several times, and this is the case in the youngest readers. Therefore, from the beginning, lexical processing is delayed because the familiarity checking stage takes more time, leading to an increased fixation duration. The delay is greater in the youngest readers than in more experienced developing readers, and generally in developing readers than in expert readers. Second, once the familiarity checking stage is completed, the full lexical access stage can start. This stage may also be affected by frequency, because of a more difficult or slower recovery of the meaning of low-frequency words than of high-frequency ones. Therefore, another issue is whether, in addition to the delay observed at this early *L1* stage, frequency still affects the *L2* stage, which specifically concerns lexical access. This interpretation is in line with the conclusions of Reichle et al. (2013) and Perfetti (2007) who posited that the major difference between children and adults mainly reflects differences in lexical and linguistic efficiency. This pattern might mean that word recognition is a key step in reading in children that does not vanish when higher processes are involved, as in text reading. The effect of lexical frequency, which is found consistently in the reading of isolated words, lasts in situations where higher-order processes are required, as in text reading. We found that frequency effects were stronger in third graders than in fifth graders. This developmental effect can be seen in purely quantitative terms, i.e. the less exposed a subject is to written words, the greater the frequency effect is. The effect of exposure might have a direct impact on the familiarity checking stage. Indeed, first fixations for frequent words were very similar in third and fifth graders ($M = 292$ ms and $M = 305$ ms, respectively), while they differed for lower frequency words ($M = 390$ ms and $M = 344$ ms, respectively).

An alternative interpretation based on qualitative changes occurring throughout reading development is possible. It is widely thought that sublexical processing based on phonological activation progressively decreases with reading development (Ziegler, Lété, & Grainger, 2014). Eye movement analysis has shown that shorter saccade amplitude in reading suggests a reliance on the sublexical phonological procedure (De Luca, Borrelli, Judica, Spinelli, & Zoccolotti, 2002), consistent with the grasp of smaller units at each fixation. In our study centered on frequency effect, spatial measures were focused on fixation and refixation probability and not on saccade amplitude. We did not find any difference in either of the developing reader groups, while we cannot exclude an impact of age on saccade amplitude. However, if our group of younger children grasped only sublexical units, no lexical frequency effect should have been observed. Therefore, we argue that between third and fifth grade, the changes that occur are mostly quantitative and reflect a progressively increasing speed in the first stage of word processing.

Considering expert readers, previous studies found a frequency effect in text reading in first fixation and gaze duration (Chamberland et al., 2013; Miellet et al., 2007; Raney & Rayner, 1996), while we did not, even though the target words all came from an adult lexical database. There could be several explanations for this discrepancy. The first concerns word predictability. It has already been shown that when target words are highly predictable, the frequency effect may disappear (Serenio, Hand, Shahid, Yao, & O'Donnell, 2017). Predictability (reported in Table 2) was significantly higher in low- than in high-frequency words in expert readers. Thus, the frequency effect usually observed in first fixation and gaze duration might have been suppressed or reduced in expert readers. A second reason concerns the frequency range that we used. For example, our low-frequency words had a mean frequency count of 20 occurrences per million, while those used by Raney and Rayner (1995, 1996) were 7 per million. Likewise, our high-frequency words had a mean frequency count

of 268 per million, while those used by Miellet et al. (2007) were 457 per million. In other words, our low-frequency words might not have been sufficiently infrequent and/or our high-frequency words might not have been sufficiently frequent to evidence an early frequency effect in expert readers. Finally, expert readers skipped more target words (both high- and low-frequency) than developing readers, so a lack of data in expert readers might have prevented the emergence of a significant early frequency effect.

6. Limitations and future remedies

Our study has some inherent limitations due to our experimental constraints. Indeed, the challenge of our study was to examine the frequency effect in developing and in expert readers at the same time. Thus, we inevitably had the issue of material selection to deal with. We chose to use the same text for developing and expert readers rather than a text adapted for each age group. The frequency of target words was controlled for each group, even though the text was more adapted to developing rather than to expert readers. In other words, the text was probably easier to read for adults than for children. The level of difficulty is known to impact eye movement behavior (Blythe & Joseph, 2011). Thus, in future research, in addition to using a common text, it would be useful to use another text for expert readers with a suitable level of difficulty. This second text, which should be matched to the first as much as possible, would offer a benchmark for the frequency effect in expert readers.

Finally, as noted at the end of the previous section, the probability of skipping is higher in expert than in developing readers. This raises the question of whether the number of target words used in our study (10 high- and 10 low-frequency words) was sufficient to examine the frequency effect, especially in expert readers.

7. Conclusion

The present study shows that with age-appropriate stimuli, a frequency effect is observed in both developing and expert readers in text reading, a situation more akin to natural reading

than the reading of independent sentences. While the frequency effect was observed in all participants on both selected temporal (and for expert readers at least in total reading time) and spatial measures, its magnitude differed according to the reading level. It was greater in third graders than in fifth graders and expert readers, and greater in fifth graders than in expert readers. Furthermore, it appeared in the early measures of fixation duration in both third graders and fifth graders, while it was apparent only in the total reading time in expert readers. An extended exposure to low-frequency words with increasing age provides readers with some of the benefits that are already observed in the early stage of lexical processing. These benefits could reverberate down to the subsequent stages of processing and facilitate them at each stage, thus allowing better lexical access and better semantic integration of a given word in the linguistic context where it appears. Further studies are needed to better understand how the impact of lexical frequency changes with proficiency.

References

Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon*, 5(3), 436-461.

- Bates, D., Maechler, M., & Dai, B. (2013). Lme4: Linear mixed-effects models using Eigen and Eigenfaces. R Package Version 0.999999-4.
- Blythe, H. I., Livversedge, S. P., Joseph, H. S., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research*, 49(12), 1583-1591.
- Blythe, H. I., & Joseph, H. (2011). Children's eye movements during reading. In S. Livversedge, I. Gilchrist, & S. Everling (Eds.). *The Oxford handbook of eye movements* (pp. 643-662). Oxford University Press.
- Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology*, 58(5), 412-424.
- Chamberland, C., Saint-Aubin J., & Lgre, M. (2013). The impact of text repetition on content and function words during reading: Further evidence from eye movements. *Canadian Journal of Experimental Psychology*, 67(2), 94-99.
- Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency Effects in Monolingual and Bilingual Natural Reading. *Psychonomic Bulletin and Review*, 22(5), 1216-1234.
- Cordier, F., Croizet, J. -C., & Rigalleau, F. (2013). Comparing nouns and verbs in a lexical task. *Journal of Psycholinguistic Research*, 42(1), 21-35.
- De Leeuw, L., Segers, E., & Verhoeven, L. (2016). The effect of student-related and text-related characteristics on student's reading behaviour and text comprehension: An eye movement study. *Scientific Studies of Reading*, 20(3), 248-263.
- De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye movement study of developmental dyslexia. *Brain and Language*, 80(3), 617-626.

- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological Review*, 112(4), 777-813.
- 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(3), 417-429.
- Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, 49(24), 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(6), 1430-1440.
- Inhoff, A. W. (1984). Two stages of word processing in the reading of prose during eye fixations in reading. *Journal of Verbal Learning & Verbal Behavior*, 23(5), 612-624.
- Joseph, H. S., & Liversedge, S. P. (2013). Children's and adults' on-line processing of syntactically ambiguous sentences during reading. *PLoS One*, 8(1), e54141.
- Joseph, H. S., Nation, K., & Liversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review*, 42(2), 207-222.
- Lefavrais, P. (1967). *Test de l'Alouette: manuel*. Paris: Les Editions du Centre de Psychologie Appliquée.
- Legge, G. E., Hooven, T. A., Klitz, T. S., Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: New insights from an ideal-observer model of reading. *Vision Research*, 42(18), 2219-2234.

- Lo, S. & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, 6, 1-16.
- Mielliet, S., Sparrow, L., & Sereno, S.C. (2007). Word frequency and predictability effects in reading French: An evaluation of the E-Z Reader model. *Psychonomic Bulletin & Review*, 14(4), 762-769.
- Miller, B. & O'Donnell, C. (2013). Opening a window into reading development: eye movements' role within a broader literacy research framework. *School Psychology Review*, 42(2), 123-139.
- New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516-524.
- Peereman, R., Lété, B., & Sprenger-Charolles L. (2007). Manulex-infra: Distributional characteristics of grapheme-phoneme mappings and infralexical and lexical units in child-directed written material. *Behavior Research Methods*, 39(3), 593-603.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357-383.
- Pollatsek, A., Reichle, E., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, 52(1), 1-56.
- R Development Core Team. (2013). R: A language and environment for statistical computing. Retrieved from <http://www.R-project.org>.
- Radach, R., Kennedy, A. (2004). Theoretical perspectives on eye movements in reading. Past controversies, current deficits and an agenda for future research. *European Journal of Cognitive Psychology*, 16(1-2), 3-26.
- Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. *Psychological Research*, 72(6), 675-688.

- Radach, R., Schmitt, C., Glover, L., & Huestegge, L. (2009). How children read for comprehension: Eye movements in developing readers. In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.) *Beyond decoding: The biological and behavioral foundations of reading comprehension* (pp. 75–106). New York: Guildford Press.
- Raney, G. E. & Rayner, K. (1995). Word frequency effects during two readings of a text. *Canadian Journal of Experimental Psychology*, 49(2), 151-172.
- Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of Visualized Experiments*, 83, 1-7.
- Rau, A. K., Moeller, K., & Landerl, K. (2014). The transition from sublexical to lexical processing in a consistent orthography: An eye-tracking study. *Scientific Studies of Reading*, 18(3), 224-233.
- Raven, J. (1998). *Les Progressive Matrices Couleurs (CPM ou PM47)*. Paris: Editions et Applications Psychologiques.
- Rayner, K. (1998). Eye movements in reading and information processing: Twenty years of research. *Psychological Bulletin*, 124(3), 372–422.
- Rayner, K. (2009). The Thirty-fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457-1506.
- Rayner, K. & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14(3), 191-201.

- Rayner, K., Juhasz, B. J., & Pollatsek, A. Eye Movements During Reading. In: Snowling, M. J., Hulme, C., (Eds.) *The Science of Reading: A Handbook*. Malden, MA: Blackwell Publishing; 2005. pp. 79-97.
- Rayner, K., & Livversedge, S. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Livversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 751-766). England: Oxford.
- Rayner, K. & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, 3(2), 245-248.
- Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, 4(3-4), 21-49.
- Rayner, K. (1998). Eye movements in reading and information processing: Twenty years of research. *Psychological Bulletin*, 124(3), 372-422.
- Reichle, E. D., Livversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H., White, S., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, 33(2), 110-149.
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Towards a model of eye movement control in reading. *Psychological Review*, 105(1), 125-157.
- Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, 26(4), 445-476.
- Russell V. Lenth. (2016). Least-Squares Means: The R Package lsmeans. *Journal of Statistical Software*, 69(1), 1-33.

- Schilling, H. E. H., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, 26(6), 1270-1281.
- Schroeder, S., Hyönä, J. & Liversedge, S. P. (2015). Developmental eye-tracking research in reading: Introduction to the Special Issue. *Journal of Cognitive Psychology*, 27(5), 500-510.
- Sereno, S. C., Hand, C. J., Shahid, A., Yao, B., & O'Donnell, P. J. (2017). Testing the limits of contextual constraint: Interactions with word frequency and parafoveal preview during fluent reading. *Quarterly Journal of Experimental Psychology*, 71(1), 302–313.
- Tiffin-Richards, S. P., & Schroeder, S. (2015). Word length and frequency effects on children's eye movements during silent reading. *Vision Research*, 113(Pt A), 33-43.
- Valle, A., Binder, K. S., Walsh, K., Nemier, C., & Bangs, K. (2013). Eye movements, prosody, and word frequency among average and high skilled second grade readers. *School Psychology Review*, 42(2), 171-190.
- Vitu, F. (2005). Visual extraction processes and regressive saccades in reading. In G. Underwood (Ed.). *Cognitive processes in eye guidance* (pp. 1–32). Oxford: Oxford University Press.
- Vorstius, C., Radach, R., & Lonigan, C. J. (2014). Eye movements in developing readers: A comparison of silent and oral sentence reading. *Visual Cognition*, 22(3), 458-485.
- Wallot, S., Hollis, G., & van Rooij, M. (2013). Connected Text Reading and Differences in Text Reading Fluency in Adult Readers, *PloS One*, 8(8), e71914.
- Wang, H. C., Pomplun, M., Ko, H. W., Chen M. L., & Rayner, K. (2010). Estimating the effect of word predictability on eye movements in Chinese reading using latent semantic analysis and transitional probability. *Quarterly Journal of Experimental Psychology*, 63(7), 1374-1386.

- Whitford, V., & Titone, D. (2012). Second-language experience modulates first- and second-language word frequency effects: Evidence from eye movement measures of natural paragraph reading. *Psychonomic Bulletin & Review*, 19(1), 73-80.
- Ziegler, J., Bertrand, D., Lété, B., & Grainger, J. (2014). Orthographic and phonological contributions to reading development: Tracking developmental trajectories using masked priming. *Developmental Psychology*, 50(4), 1026-1036.