Routledge
Taylor & Francis Group

Eye movements in developing readers: A comparison of silent and oral sentence reading

Christian Vorstius[1]*, Ralph Radach[1], and Christopher J. Lonigan[2]

[1]University of Wuppertal, Wuppertal, North Rhine-Westphalia, Germany
[2]Florida State University, Tallahassee, Florida, USA

We present sentence reading data from a large-scale study with children ($N = 632$), focusing on three key research questions. (1) What are the trajectories of reading development in oral as compared to silent reading? (2) How are word frequency effects developing and are changes differentially affected by reading mode? (3) Are there systematic differences between better and weaker comprehenders when reading silently vs. aloud? Results illuminate a number of differences between reading modes, including more and prolonged fixations in oral reading, along with less inter-word regressions and attenuated effects of word frequency. Weaker comprehenders were slower, especially in oral reading and showed less flexibility in the allocation of word processing time. Differences between reading modes can be explained by additional processing demands imposed by concurrent articulation and eye–voice coordination when reading aloud.

INTRODUCTION

Over the last three decades, our understanding of the perceptual and cognitive processes underlying skilled reading has grown substantially and visuomotor

research has contributed significantly to this progress (see Radach & Kennedy, 2004, 2013; Rayner, 1998, 2009 for reviews). A number of well-developed quantitative models are available (Engbert, Nuthmann, Richter, & Kliegl, 2005; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reilly & Radach, 2006; Yang & McConkie, 2001) allowing accurate modelling of eye movements and contributing to the enhancement of our understanding of linguistic, visual, attentional, and oculomotor control mechanisms and their complex interplay.

Notwithstanding these advances, surprisingly little work has focused on children and the developmental aspects of reading. Compared to the substantial literature on adult reading, the number of publications on eye movements in developing readers is quite limited (see Blythe & Joseph, 2011; Radach, Schmitten, Glover, & Huestegge, 2008, for reviews). Especially noteworthy are a number of pioneering studies that laid much of the groundwork for today's more elaborate research (e.g., Buswell, 1922; Taylor, Frackenpohl, & Pettee, 1960; McConkie et al., 1991; Rayner, 1986). These early studies examined elementary school students across various grades, with quantitative analyses mostly restricted to global parameters such as average fixation duration, number of fixations per 100 words, and overall regression frequency (see Table 20.1 in McConkie et al., 1991, for details). Not surprisingly, this work documented a steady reduction of fixation durations and number of fixations from grade to grade, whereas the decrease was much less pronounced for the proportion of regressive saccades (eye movements against the reading direction). It is noteworthy that very similar general trends appear in all four studies despite various differences in tracking technology, reading materials (sentences vs. paragraphs), and reading mode (silent vs. aloud).

McConkie et al. (1991) opened a new chapter in the analysis of children's eye movements, introducing novel issues and research tools to the topic, such as frequency distributions of viewing time measures (see also, McConkie & Dyre, 2000), fine grain examinations of saccade landing sites within words (McConkie, Kerr, Reddix, & Zola, 1988; McConkie, Kerr, Reddix, Zola, & Jacobs, 1989), and the first quantitative analyses of relations between eye movement parameters and psychometric reading assessments. This was also one of the first studies in which the now common dissociation of viewing times into initial fixation durations, gaze duration and total viewing time (Inhoff & Radach, 1998; Rayner, 1998) was applied to research on developing readers. The idea of this decomposition is to delineate the time course of word processing into time intervals that reflect early orthographic and lexical processing (initial fixations), full lexical access (refixation times) and time spent on integration of word meanings into representations on the sentence and text level (rereading times). The analyses of word viewing times in the present paper will follow this logic (see Radach & Kennedy, 2004, for a discussion).

A prime example demonstrating how critical information can be overlooked when relying solely on global measures is the case of regression rates. The

overall proportion of regressions (leftward saccades) seems to stay fairly constant across readers from different grades. However, when analysing the data on a word-based level it becomes apparent that intraword regressions (regressive refixations within the same word) become less frequent, whereas interword regressions (saccades going back to text left of the currently fixated word) become more and more frequent as reading skill develops (McConkie et al., 1991). Intra- and interword regressions reflect qualitatively different cognitive processes in adult reading, with many regressions executed in the interest of high level comprehension (e.g., Inhoff, Weger, & Radach, 2005). Assuming that this is also the case for beginning readers, the age-related increase in interword regressions may reflect a qualitative change from a dominance of decoding-related effort towards allocating more and more cognitive resources to comprehension.

Recent studies examining eye movement patterns in children have used experimental designs to examine specific research questions in specific age ranges. These questions include differences between children and adults in perceptual span size (Häikiö, Bertram, Hyönä, & Niemi, 2009; Rayner, 1986), word length and frequency effects (Blythe, Liversedge, Joseph, White, & Rayner, 2009; Blythe, Häkiö, Bertram, Liversedge, & Hyönä, 2011; Huestegge, Radach, Corbic, & Huestegge, 2009; Hyönä & Olson, 1995; Joseph, Liversedge, Blythe, White, & Rayner, 2009; Joseph, Nation, & Liversedge, 2013), and binocular coordination (Blythe, Liversedge, Joseph, White, Findaly, & Rayner, 2006). Very few studies have so far focused on higher level processing beyond the word level (e.g., Joseph et al., 2008; Joseph & Liversedge, 2013; Vorstius, Radach, Mayer, & Lonigan, 2013). Taken together, data collected in this work suggest that with development, the spatial extent of the perceptual span increases, word viewing times decrease, fewer fixations and regressions are made, and fixation probabilities and refixation frequencies are reduced. Manipulations of word processing difficulty are generally reflected in children's eye movements, although often such effects are less clear cut compared to skilled reading in adults (see Blythe & Joseph, 2011, for a comprehensive review).

The present work is aimed at further advancing our understanding of reading development using a combination of experimental and corpus based methodology. Similar to the materials developed by Schilling, Rayner, and Chumbley (1998) and Kliegl, Grabner, Rolfs, and Engbert (2004), participants were asked to read a sample of single, unconnected sentences. These sentences can be analysed as a textual corpus similar to the English (Vitu & McConkie, 2000) and German (Radach & McConkie, 1998) Gulliver's Travels corpora, and the English/French newspaper corpus (Pynte & Kennedy, 2006). However, due to the inclusion and variation of well-controlled target words, our materials also allow for factorial analyses strategies (see Radach, Huestegge, & Reilly, 2008, for a similar approach).

To date, no direct comparison of eye movements in oral vs. silent reading in children has been published. However, such a comparison is interesting from two different, but related perspectives. First, from a theoretical point of view, we

know little about how, in comparison to silent reading, the more complex process of reading aloud is controlled. This includes the question of how the addition of a language production component in oral reading influences the ongoing processes of spatially distributed information acquisition, word recognition and comprehension. A related topic concerns the coordination of both processing streams, receptive and productive, as evident in the eye–voice span (Inhoff, Solomon, Radach, & Seymour, 2011). Second, from an educational point of view, oral reading is the starting point of development and the focus of typical early reading curricula. Reading aloud provides an easy way for teachers to track errors and often oral reading performance is assumed to provide a valid and reliable ad hoc assessment for the general reading ability of a child. In our view, it is not quite clear whether such a generalization is justified, as relatively little is known about the precise nature of the differences between silent and oral reading, especially regarding the dynamics of development in both modes of reading.

Relating reading mode and comprehension, it appears quite reasonable to assume that reading aloud enhances comprehension because readers can benefit from additional auditory input as they are monitoring their own articulation (Kragler, 1995). Moreover, the ongoing process of grapheme–phoneme conversion might draw more resources towards sublexical processing units, which may benefit both word recognition and comprehension (Kragler, 1995; Miller & Smith, 1990; Prior & Welling, 2001; Swalm, 1973).

The alternative view, that reading aloud may hinder comprehension goes back to Jones and Lockhart (1919), who first suggested that readers must devote extra resources to pronunciation, intonation, monitoring, etc., leaving less capacity for comprehension. As one specific variant of this view, it has been suggested that processing of a word may come to an end after its pronunciation when reading aloud. In this case the focus on the grapheme-to-sound route may actually impede access to the more direct lexical and/or semantic information (Juel & Holmes, 1981). These authors also suggested that the effect of reading mode on comprehension may depend on the state of development, with beginning and struggling readers showing better achievement during oral reading, readers in higher grades (3rd and 4th) showing no difference and skilled (adult) readers exhibiting better results during silent reading. The question of whether such a developmental trajectory is reflected in eye movement parameters has yet to be answered.

The present study is part of a large-scale research project. The general design combines cross-sectional and longitudinal comparisons, but, as a first step, the present paper reports cross-sectional data of grades 1 to 5, collected during year one of the project. Due to space limitations we decided to focus on analyses of viewing times (although selected spatial parameters will be reported), examining three main research questions: (1) What are the trajectories of reading development in oral as compared to silent reading? (2) Using word frequency as a standard way to manipulate word processing difficulty, we raise the question

of how frequency effects change over the course of reading development and how this change is differentially affected by reading mode. (3) Finally, we explore the effects of comprehension level on eye movements, asking whether there are systematic differences between stronger and weaker comprehenders in the processing of words when reading silently vs. aloud.

METHODS

Participants

Participants were part of a large-scale study on reading development in which 1906 children from pre-k through 5th grade were tested in year one. Out of this sample, 632 children in grades 1–5 were asked to read a sentence both silently and aloud while their eye movements were being monitored ($n = 153$, $n = 117$, $n = 134$, $n = 111$, and $n = 117$, respectively, from grade 1 to 5). Participating students attended classrooms in different schools in North Florida and were ethnically and economically diverse. Written informed consent was obtained from a parent or legal guardian and child assent was collected at the beginning of the testing session. Table 1 presents key demographic information.

Materials

Participants were asked to read simple declarative sentences, including 3 practice and 48 experimental sentences in each reading mode (102 trials total). Experimental sentences consisted of 9–14 words and had a mean length of 62.6 ($SD = 3.1$) letters in the silent condition and 62.9 ($SD = 3.5$) letters in the aloud condition. Since the data collected in this study are intended to be combined with other assessments of reading for analyses of individual

TABLE 1
Key demographics for analysed data sets

		Grade				
		1	2	3	4	5
Age (year; month)		7;0	7;11	9;0	9;11	10;10
Gender	Male	47.2%	36.5%	53%	51.4%	42.5%
	Female	51.2%	61.7%	47%	47.7%	55.8%
Race	White	64.2%	71.3%	64.2%	65.8%	61.9%
	Black / African American	20.3%	12.2%	21.6%	18.0%	24.8%
	Asian	0.8%	2.2%	2.2%	4.5%	3.5%
	Others	4.9%	4.3%	1.5%	–	3.5%

1st grade, $n = 123$; 2nd grade, $n = 115$; 3rd grade, $n = 134$; 4th grade, $n = 111$; 5th grade, $n = 113$; $N = 596$. Some participants did not provide all demographics, therefore not all categories add up to 100%.

differences, sentences were assigned to the two reading mode conditions in a fixed way so that all participants read identical sentences in the same order within the respective blocks. One target word was embedded in each sentence. Target words were always nouns, 4–7 characters long (25% in each length) and either of high or low word frequency (50% in each word length). The U-measure of word frequencies based on grade-appropriate text samples developed by Zeno, Ivens, Millard, and Duvvuri (1995) was used to ensure that target words were part of common vocabulary and reading instruction for young readers, while maximizing the contrasts for low vs. high frequency at each grade level. The resulting mean word frequencies amounted to 8.80 and 99.67 for low vs. high frequency words, corresponding to 4.47 and 80.54 words per million in the CELEX word corpus (Baayen, Piepenbrock, & van Rijn, 1993), with a correlation of $r = .89$ between both frequency measures. The grade-specific word frequency measures provided by Zeno et al. varied between 17.94 vs. 191.98 for grade 1 and 9.58 vs. 111.60 for grade 5, indicating that the degree of familiarity with low vs. high frequency target words was similar across grades. Target words were also matched for number of morphemes (both 1.08 for low and high frequency words) and syllables, with means of 1.67 syllables for low frequency words and 1.58 syllables for high frequency target words. Pre- and post-target words were also controlled for word length, with the length of the preceding word amounting to 5.35 letters of low frequency and 5.52 letters for high frequency targets, and the length of the following words being 3.85 letters for low frequency and 4.00 letters for high frequency target words (see Krügel & Engbert, 2010; Radach & Kempe, 1993, for effects of neighbouring words on eye movements). The following examples illustrate the nature of the sentences used in this study:

(1) `One time mother found a dirty sock right under Linda's bed.` (sock: low frequency)
(2) `Paul and grandpa brought a large fish back from the market today.` (fish: high frequency)

Comprehension questions appeared at fixed random intervals after one out of four sentences and were always related to the preceding sentence. Questions targeted both simple (e.g., location, object) and complex (e.g., causality, inference) semantic relations so that both reading for surface features and deep meaning was encouraged. For example, the comprehension question for (1) was, "What was laying under the bed?", and for (2) the question, "What is grandpa going to cook?", was used. Questions were answered orally and the experimenter recorded given answers on the participant log sheet. Answers were scored with 0 points for missing or incorrect answers, 1 point for partially correct answers, and 2 points for correct answers, resulting in a possible maximum comprehension score of 24 points.

Apparatus

Single line sentences were displayed one-by-one in black on a light gray background using a 21-inch flat-panel monitor. Display resolution was set to 1024×768 pixels with a refresh rate of 120 Hz. Texts were presented in Courier New font 15pt, left aligned and vertically centred on the screen. Viewing distance between each reader's eyes and the monitor was adjusted to 68 cm. At this distance, each letter subtended 0.33° of visual angle laterally. Viewing was binocular and eye movements from both eyes were recorded using the EyeLink1000, sampling at 500Hz. A three-point calibration was performed at the beginning of each task and after every question (on average every six trials). Mean average position error in an accuracy validation routine was not to exceed 0.33° of visual angle. A drift check before every trial ensured accuracy between calibrations. If the drift check showed a deviation from more than 0.33° of visual angle, an additional calibration was performed. The online saccade parser of the eye tracking system was set to detect saccades with an amplitude of 0.15° or greater, using an acceleration threshold of 8000°/sec^2 and a velocity threshold of 30°/sec. These settings have proven to produce highly accurate and reliable data in multiple reading studies across different laboratories (see Inhoff & Radach, 1998; McConkie, 1981, for detailed discussions of methodological issues).

A basic vision screening was performed using an Optec 5000 vision tester. Tests included near and far visual acuity, lateral phoria, vertical phoria, and a fusion test. Results were reported to parents if a participant failed one or multiple tests. However, children were only excluded from participation based on the vision screening if their (corrected) acuity did not allow them to read the text presented on the screen during the experiment.

Procedure

Experimenters went to the student's classroom and escorted him/her to the designated testing room in the school. After obtaining the child's assent the basic vision screening was performed. Afterwards, participants were seated in front of the presentation monitor and received the directions for the upcoming task on the display screen in front of them. For children in lower grades, experimenters read the directions out loud to the children to ensure comprehension. Participants received identical directions for all reading tasks, instructing them to read every text so that they understand its meaning and are able to answer comprehension questions. The silent and oral sentence reading tasks were two of a total of five reading tasks administered to all participating children. Silent sentence reading was the first task administered, and oral reading the third or fourth task in this sequence, depending somewhat on reading speed and scheduling constraints. The other tasks consisted of a silent paragraph reading task and two Landolt scanning tasks (Hillen et al., 2013). There were short breaks between the tasks

and a testing session usually lasted about one school hour (45 min) including the transfer back and forth from the classroom. For most children, testing was completed within two sessions. Within each session, children could take breaks between tasks or before calibrations, if necessary.

Data analysis

Raw eye movement data were converted to pairs of fixations and incoming saccades and then aggregated to fixation-based and word-based matrices using the software suite EyeMap (Tang, Reilly, & Vorstius, 2012) as well as SPSS. During this procedure all scan paths were also inspected visually to detect any possible problems. Data analyses reported are based on right eye fixations and saccades.

Thirty-six data sets had to be excluded, as participants did not follow instructions, moved the head extensively during the task, or were not able to read (1st grade: 30, 2nd grade: 2, and 5th grade: 4). Practice trials, first and last word on each line, and any fixations with blinks or durations shorter than 70 ms or longer than 1200 ms, gaze durations longer than 2400 ms or total viewing times longer than 4800 ms were excluded from the analyses (∼6% in each task). No further data cleaning was performed.

Inferential statistics for continuous variables are based on linear mixed models (LMMs) specifying subjects (= participants) and items (= words) as random effects (Baayen, Davidson, & Bates, 2008; Kliegl, Masson, & Richter, 2010). Analyses for all words included Grade (1st, 2nd, 3rd, 4th, 5th) as a between subject factor and Reading Mode (silent vs. aloud) as a within subject factor, both as fixed effects. For target words analyses, Word Frequency (high vs. low) was added as an additional fixed effect to the model. For analyses with respect to comprehension, participants were divided into stronger and weaker comprehenders based on their answers to the comprehension questions. We used median-split by grade to create the groups. Values of temporal parameters were log-transformed for statistical analyses to better fit normal distributions. Figures and tables presented use non-transformed values for ease of interpretation.

All effects were estimated using the lmer programme from the lme4 package (Bates, Mächler, & Bolker, 2012) in the R environment for statistical computing (version 3.0.1; R Core Team, 2013). Regression coefficients, standard errors (*SE*), and *t*-values (*z*-values for tests with binary, which were analysed with models using the binomial distribution with a logit link function) are reported. For all tests, we used the two-tailed criterion ($t > |2|$ or $z >= |2|$), corresponding to a 5% error criterion for significance.

Contrasts were defined using successive differences coding (contr.sdif from the MASS package, Venables & Ripley, 2002). In this coding, contrast coefficients are chosen so that the coded coefficients in a one-way layout are the

differences between the means of the second and first levels, the third and second levels, and so on. All 2-way interactions are coded accordingly.

Dependent variables

Key temporal parameters used in the analyses are initial fixation durations, refixation time, and rereading time, providing a full partitioning of the total time spent fixating on every word (Radach & Kennedy, 2004). Initial fixation duration is defined as the duration of the first fixation on the word, irrespective of whether or not the word was subsequently fixated again. Refixation time is the time spent viewing a word during first pass reading (i.e., before leaving the word for the first time) minus the initial fixation duration. Rereading time is the time spent on a word after first pass reading. For refixation and rereading times intraword saccade durations are excluded. Note that the sum of initial fixation durations and refixation times per word corresponds to the widely used measure of gaze duration or first pass reading time. In addition, total viewing time is reported for some analyses. Total viewing time is defined as the summed duration of all fixations on a word (initial fixation + refixation time + rereading time).

Other key variables include the proportions of refixations and regressions. The proportion of refixations reflects the number of fixations within the first reading pass that, out of all fixations, land on the same word as the previous fixation (and thus constitute a refixation on that word). The proportion of regressions was calculated for interword saccades and quantifies the amount of fixations that were preceded by a leftward saccade (against the reading direction).

In addition, spatial parameters include saccade amplitudes, initial progressive saccade landing positions, and saccade launch distances. Saccade amplitudes describe the extent of the incoming saccade in letter units. Initial landing positions refer to the letter position within the word on which the initial fixation in that word landed (Rayner, 1979). The space preceding the word is coded as letter 0. Launch distance is the distance between the word beginning of the currently fixated word and the starting position of the incoming (preceding) saccade (see Inhoff & Radach, 1998; Inhoff & Weger, 2003; Rayner, 1998, for discussions of oculomotor measures). Spatial parameters were calculated for 5-letter words only and are based exclusively on incoming progressive saccades.

RESULTS

Analyses for all words in the sentence corpus

Initial fixation duration.   Mean initial fixation durations decreased across grades (Figure 1a) with significant differences between 1st and 2nd ($b = -0.114$, $SE = 0.016$, $t = -7.0$) and 3rd and 4th grade ($b = -0.073$, $SE = 0.016$, $t = -4.5$).
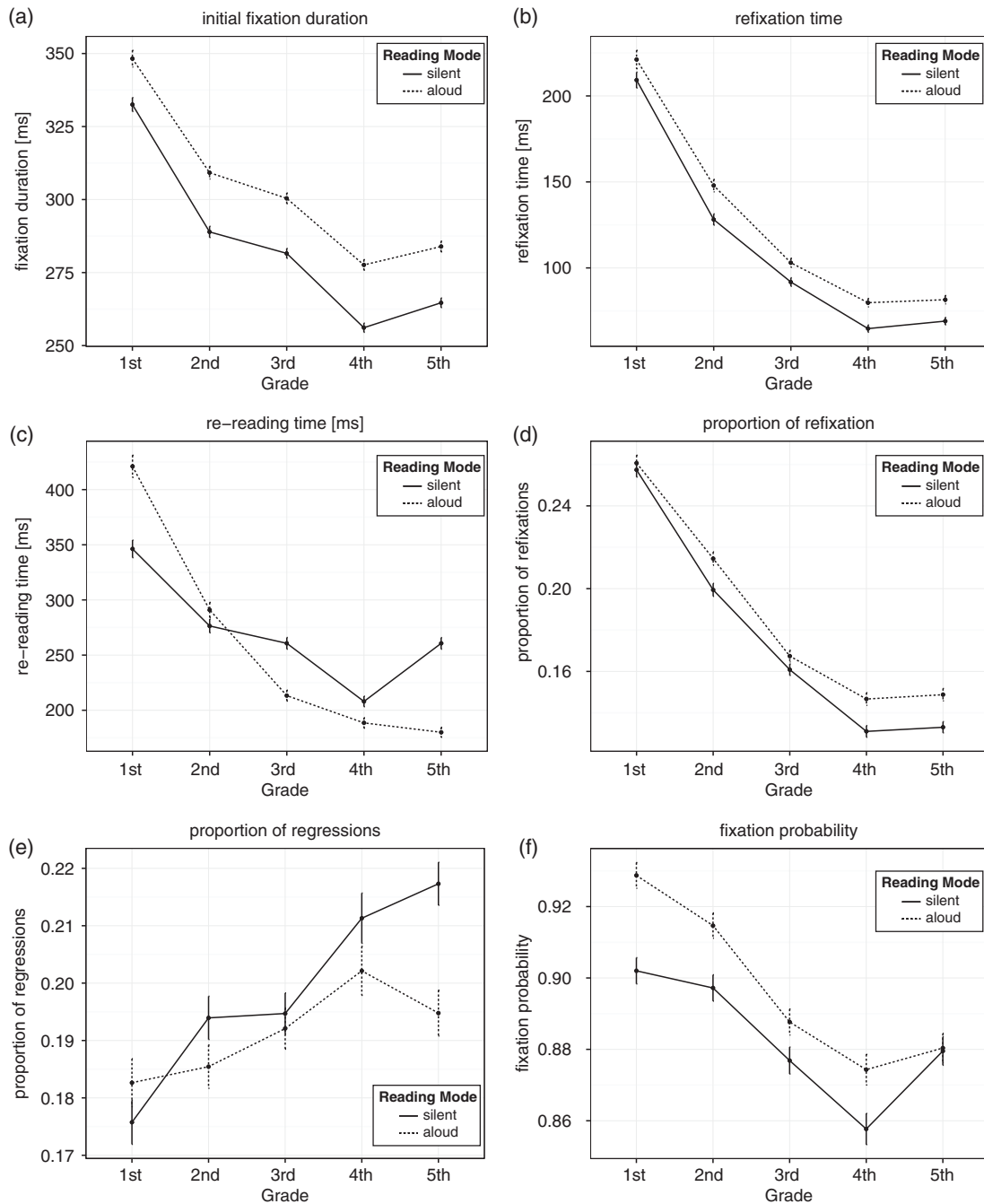
Figure 1.   Key parameters by Grade and Reading Mode for all words. Temporal measures (a–d) are given in ms (*M* and *SE*), (e) and (f) depict proportions of refixations and regressions, respectively.

In the reading aloud task, initial fixation durations were significantly longer, compared to the silent reading mode ($b = 0.057$, $SE = 0.002$, $t = 32.9$). In addition, there was a significant interaction for Grade and Reading Mode between 1st and 2nd grade ($b = 0.028$, $SE = 0.004$, $t = 6.5$), 3rd and 4th grade ($b = 0.014$, $SE = 0.004$, $t = 3.5$), as well as 4th and 5th grade ($b = -0.013$, $SE = 0.004$, $t = -3.1$), indicating that the difference between reading modes was larger in

TABLE 2
Means (*M*) in ms and Standard Error (*SE*) of key temporal parameters for all words across grades for silent and aloud reading modes

			Grade				
			1	2	3	4	5
Inital fixation duration	Silent	M	333	289	282	256	265
		SE	1.2	1.0	0.9	0.8	0.8
	Aloud	M	348	309	300	278	284
		SE	1.4	1.1	0.9	0.9	0.9
Refixation time	Silent	M	209	128	92	65	69
		SE	2.3	1.6	1.2	1.1	1.1
	Aloud	M	221	148	103	80	82
		SE	2.7	1.8	1.3	1.2	1.2
Rereading time	Silent	M	346	276	261	208	261
		SE	4.0	3.1	2.6	2.4	2.6
	Aloud	M	421	291	213	189	180
		SE	5.0	3.4	2.5	2.4	2.2

second compared to first grade and smaller in 5th than in 4th grade (see Table 2). No other interactions were significant (all ts < 1). Similar patterns were found for 1st of 2 and 2nd of 2 fixation durations, indicating that fixation durations were generally longer when reading aloud, independent of their functional role.

Refixation time.   Refixation time also decreased across grades. Significant differences emerged between 1st and 2nd grade ($b = -0.563$, $SE = 0.068$, $t = -8.23$), 2nd and 3rd grade ($b = -0.428$, $SE = 0.067$, $t = -6.4$), and 3rd and 4th grade ($b = -0.259$, $SE = 0.068$, $t = -3.82$). Reading Mode also significantly affected refixation times ($b = 0.075$, $SE = 0.010$, $t = 7.49$), with silent reading having ~11 ms shorter refixation times compared to reading aloud. Significant 2-way interactions for Grade and Reading Mode were found between 1st and 2nd ($b = 0.143$, $SE = 0.025$, $t = 5.83$), 2nd and 3rd ($b = -0.049$, $SE = 0.023$, $t = -2.16$) and 3rd and 4th grade ($b = 0.066$, $SE = 0.023$, $t = 2.86$). Differences between reading mode are larger in 2nd compared to 1st grade and shorter in 3rd grade compared to and 2nd and 4th grade (see Figure 1b).

Rereading time.   Rereading times also declined across grades. However, significant differences were found only between 1st and 2nd grade ($b = -0.357$, $SE = 0.103$, $t = -3.46$). The difference between reading modes was also significant ($b = -0.225$, $SE = 0.012$, $t = -19.16$), with rereading times that were ~17 ms longer for the silent condition. In addition, all 2-way interactions for Grade and Reading Mode were significant (see Figure 1c). Interestingly, the difference between reading silently and aloud almost disappears from 1st to 2nd

TABLE 3

Mean proportion (*M*) and Standard Error (*SE*) for refixations and interword regressions for all words across grades for silent and aloud reading modes

					Grade		
			1	*2*	*3*	*4*	*5*
Proportion of refixations	Silent	*M*	.26	.20	.16	.13	.13
		SE	.002	.002	.001	.001	.001
	Aloud	*M*	.26	.21	.17	.15	.15
		SE	.002	.002	.001	.002	.001
Proportion of regressions	Silent	*M*	.18	.19	.19	.21	.22
		SE	.002	.002	.002	.002	.002
	Aloud	*M*	.18	.19	.19	.20	.19
		SE	.002	.002	.002	.002	.002
Fixation probability	Silent	*M*	.902	.897	.877	.858	.880
		SE	.002	.002	.002	.002	.002
	Aloud	*M*	.929	.915	.889	.874	.880
		SE	.002	.002	.002	.002	.002

($b = -0.434$, $SE = 0.029$, $t = -15.04$) and flips to longer rereading times from 2nd to 3rd grade ($b = -0.383$, $SE = 0.026$, $t = -14.48$). In 4th grade the difference is smaller compared to 3rd ($b = 0.181$, $SE = 0.027$, $t = 6.64$) and 5th grade ($b = -0.393$, $SE = 0.028$, $t = -13.94$).

Proportion of refixations.   The proportion of refixations (among all fixations) decreased across grades with significant differences between 1st and 2nd ($b = -0.053$, $SE = 0.007$, $t = -7.47$), 2nd and 3rd ($b = -0.045$, $SE = 0.007$, $t = -6.50$), as well as 3rd and 4th grade ($b = -0.026$, $SE = 0.007$, $t = -3.66$). There were also more refixations when reading aloud ($b = 0.013$, $SE = 0.001$, $t = 17.94$). This effect of Reading Mode was affected by grade, with significant 2-way interactions between 1st and 2nd ($b = 0.012$, $SE = 0.002$, $t = 5.07$), 2nd and 3rd ($b = -0.006$, $SE = 0.002$, $t = -3.03$), and 3rd and 4th grade ($b = 0.006$, $SE = 0.002$, $t = 2.87$). Again, the differences between reading modes emerge in 2nd grade, are somewhat attenuated in 3rd grade and more pronounced in 4th grade (see Table 3, Figure 1d).

Proportion of inter-word regressions.   Despite a small but steady numerical increase, there were no main effects for the proportion of regressions out of the currently fixated word between grades. However, Reading Mode had a significant effect in that more regressions were found in the silent reading mode ($b = -0.008$, $SE = 0.001$, $t = -6.83$). There were also significant interactions for Grade and Reading Mode between 1st and 2nd ($b = -0.020$, $SE = 0.003$, $t = -6.64$), 3rd and 4th ($b = -0.006$, $SE = 0.003$, $t = -2.13$), as well as 4th and 5th grade ($b = -0.008$, $SE = 0.003$, $t = -2.9$). In 1st grade more inter-word

regressions were made when reading aloud, whereas from 2nd grade on the proportion of regressions was higher when reading silently. This effect is more pronounced in 4th compared to 3rd grade and even more prominent in 5th grade (see Table 3, Figure 1e).

Fixation probability. The probability of fixating a word decreases across grades with significant difference between 2nd and 3rd ($b = -0.352$, $SE = 0.081$, $z = -4.32$) and 3rd and 4th grade ($b = -0.207$, $SE = 0.082$, $z = -2.53$). There was a significant main effect for Reading Mode ($b = 0.102$, $SE = 0.012$, $z = 8.81$), with a higher fixation probability when reading aloud. Interactions between Grade and Reading Mode indicate that differences between Reading Modes were more pronounced in 1st compared to 2nd ($b = -0.198$, $SE = 0.037$, $z = -5.34$), 2nd compared to 3rd grade ($b = -0.099$, $SE = 0.031$, $z = -3.19$), and 4th compared to 5th grade ($b = -0.187$, $SE = 0.030$, $z = -6.21$). In 5th grade fixation probabilities increased for both Reading Modes and this effect was more pronounced when reading silently (see Figure 1f, Table 3).

Spatial parameters. As spatial parameters are affected by word length, exemplary results are reported for 5-letter words. Similar results were obtained for other word lengths. Mean initial landing positions moved further into the word with grade. Significant effects were found from 1st to 2nd ($b = 0.191$, $SE = 0.039$, $t = 4.83$), 2nd to 3rd grade ($b = 0.153$, $SE = 0.038$, $t = 4.04$), and 3rd to 4th grade ($b = 0.101$, $SE = 0.038$, $t = 2.62$). There was no main effect on landing positions for Reading Mode, but one significant Grade by Reading Mode interaction between 1st and 2nd grade ($b = -0.075$, $SE = 0.036$, $t = -2.09$; see Table 4). Landing positions were slightly further into the word when reading aloud in 1st grade, but not different (or further into the word for silent reading) in 2nd grade (see Figure 2a).

Saccade amplitudes show the matching pattern with longer amplitudes in higher grades (1st to 2nd: $b = 0.552$, $SE = 0.117$, $t = 4.74$; 2nd to 3rd: $b = 0.413$, $SE = 0.113$, $t = 3.64$; 3rd to 4th: $b = 0.356$, $SE = 0.115$, $t = 3.10$), but there was a significant main effect for Reading Mode ($b = -0.171$, $SE = 0.025$, $t = -6.94$) with longer saccade amplitude in silent reading. In addition, there were significant Grade by Reading Mode interaction between 1st and 2nd grade ($b = -0.250$, $SE = 0.055$, $t = -4.53$) and 3rd and 4th grade ($b = -0.204$, $SE = 0.050$, $t = -4.10$). Amplitudes were larger when reading silently in 2nd compared to 1st grade and 4th compared to 3rd grade (see Figure 2b, Table 4).

Matching the pattern of initial landing position and saccade amplitude, launch distance showed significant decreases between 1st and 2nd ($b = -0.351$, $SE = 0.088$, $t = -3.98$), 2nd and 3rd ($b = -0.258$, $SE = 0.086$, $t = -3.01$), and 3rd and 4th grade ($b = -0.254$, $SE = 0.086$, $t = -2.94$), as well as a significant main effect for Reading Mode ($b = -0.158$, $SE = 0.025$, $t = 6.32$) with closer launch distances when reading aloud. In addition, significant interactions for Grade and Reading Mode were found between 1st and 2nd ($b = 0.176$, $SE = 0.056$, $t = 3.16$)

TABLE 4
Means (*M*) and Standard Error (*SE*) of key spatial parameters for 5-letter words across grades for silent and aloud reading modes

					Grade		
			1	2	3	4	5
Initial landing position	Silent	M	2.03	2.24	2.40	2.52	2.49
		SE	0.02	0.02	0.02	0.02	0.02
	Aloud	M	2.09	2.23	2.35	2.43	2.43
		SE	0.03	0.02	0.02	0.02	0.02
Saccade amplitude	Silent	M	5.52	6.19	6.53	7.01	6.89
		SE	0.04	0.04	0.03	0.04	0.04
	Aloud	M	5.49	5.93	6.31	6.58	6.44
		SE	0.04	0.03	0.03	0.04	0.03
Launch distance	Silent	M	−3.01	−3.47	−3.64	−4.01	−3.92
		SE	0.04	0.04	0.04	0.04	0.04
	Aloud	M	−2.92	−3.21	−3.68	−3.65	−3.52
		SE	0.04	0.03	0.03	0.04	0.04

Values are given in letter units; negative values for launch distance represent locations to the left in relation to the word beginning.

and 3rd and 4th grade ($b = 0.177$, $SE = 0.050$, $t = 3.52$). Differences between reading modes were greater for 2nd compared to 1st grade and 4th compared to 3rd grade (see Figure 2c, Table 4).

Overall, spatial data showed a shift of the initial landing position further into the word for higher grades. This was the result of longer saccade amplitudes that even compensated for longer launch distances. With respect to reading mode, saccade amplitudes were longer when reading silently, but neither initial landing positions nor launch distances showed significant effects, although numerical trends point into the expected direction.

Analyses of Word Frequency effects based on target words

For reasons of readability, the presentation of formal statistics for target word results is limited to Reading Mode main effects, as well as Word Frequency main effects and interaction. Descriptive statistics for all parameters are presented in Tables 5 and 6.

Initial fixation duration.   In addition to a main effect for Reading Mode ($b = 0.024$, $SE = 0.012$, $t = 2.0$), the main effect for Word Frequency was also significant ($b = 0.037$, $SE = 0.012$, $t = 3.2$). Initial fixation durations were longer for low frequency words. A Grade by Word Frequency interaction between 2nd and 3rd grade ($b = 0.032$, $SE = 0.012$, $t = 2.6$) indicates that the frequency effect
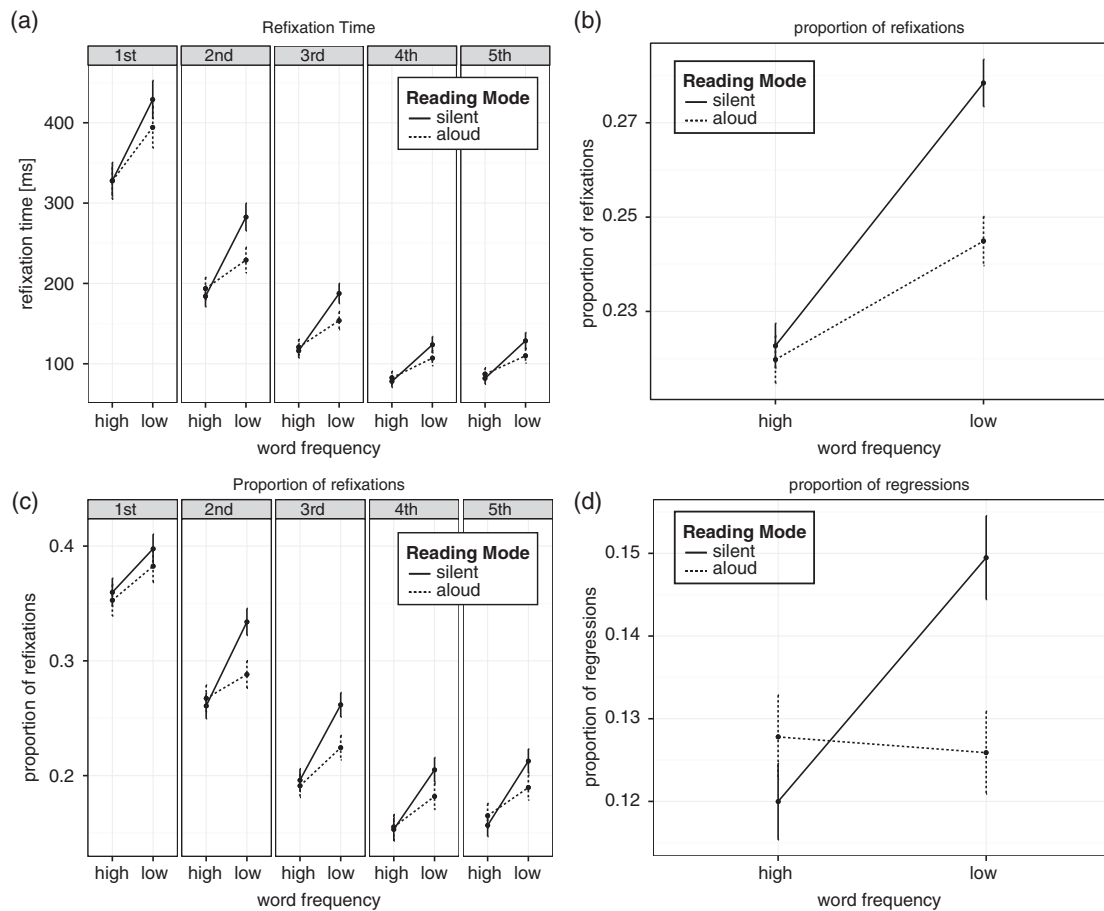
Figure 2. Depiction of interaction effects related to word frequency. (a) Interaction in Refixation Time (*M* and *SE*) between Reading Mode, Word Frequency, and Grade (1st to 2nd). Note that the overall 2-way interaction between Reading Mode and Word Frequency was also significant. (b) Reading Mode by Word Frequency interaction for proportion of refixations (*M* and *SE*). (c) Grade (1st to 2nd) by Reading Mode by Word Frequency interaction for proportion of refixations (*M* and *SE*). (d) Reading Mode by Word Frequency interaction for proportion of regressions (*M* and *SE*).

was more pronounced in 3rd compared to 2nd grade. No other effects with respect to word frequency were significant[1].

Refixation duration.   The main effects for Reading Mode were not significant. However, there was a significant main effect for Word Frequency ($b = 0.432$, $SE = 0.119$, $t = 3.63$) with longer refixation durations for low frequency words. The only significant interaction was a 3-way interaction between Grade, Reading Mode, and Word Frequency between 1st and second grade ($b = -0.412$,

---

[1] Supplementary analyses were carried out for single fixation duration, first of two and second of two fixation durations. Main results confirm the trend found initial fixation duration with shorter times from lower to higher grades and longer fixation durations when reading aloud and on low frequency target words. The Reading Mode by Word Frequency interaction was also significant as effects of word frequency were more pronounced when reading silently.

TABLE 5
Means in ms (*M*) and Standard Error (*SE*) of key temporal parameters for high and low frequency target words across grades for silent and aloud reading modes

		Word frequency		Grade				
				1	*2*	*3*	*4*	*5*
Initial fixation duration	Silent	High	M	344	307	293	270	279
			SE	5	3	3	3	3
		Low	M	363	321	322	290	301
			SE	5	4	4	3	3
	Aloud	High	M	354	319	309	287	287
			SE	5	4	3	3	3
		Low	M	368	326	322	296	288
			SE	6	4	3	3	3
Refxation time	Silent	High	M	328	184	116	78	82
			SE	10	7	4	4	4
		Low	M	429	283	188	124	129
			SE	12	9	6	5	5
	Aloud	High	M	328	194	121	83	87
			SE	11	7	5	4	4
		Low	M	394	229	154	107	110
			SE	13	8	6	5	5
Rereading time	Silent	High	M	387	282	258	197	252
			SE	16	11	9	8	9
		Low	M	461	371	334	269	332
			SE	18	15	11	10	11
	Aloud	High	M	480	314	193	165	154
			SE	20	14	8	8	7
		Low	M	568	312	210	167	162
			SE	23	13	10	9	8

$SE = 0.161$, $t = -2.56$). In 2nd grade the effect of Reading Mode is more pronounced, however, only for low frequency words (see Figure 2a).

Rereading time. A significant main effect for Reading Mode was found ($b = 0.434$, $SE = 0.129$, $t = -3.41$), but there were no significant main effects or interaction with respect to word frequency on rereading times.

Proportion of refixations. The proportion of refixations was significantly higher on low frequency words ($b = 0.291$, $SE = 0.019$, $t = 14.737$). In addition, the Reading Mode by Word Frequency interaction was significant ($b = -0.213$, $SE = 0.039$, $t = -5.39$) because the frequency effect was stronger in silent reading (Figure 2b). Finally, there was a significant Grade by Reading Mode by Word Frequency interaction between 1st and 2nd grade ($b = -0.281$, $SE = 0.127$, $t = -2.211$). The stronger effect of word frequency in silent reading only

TABLE 6
Mean proportion (*M*) and Standard Error (*SE*) for refixations and regressions for target words across grades for high and low frequency words in silent and aloud reading modes

				Grade				
				1	*2*	*3*	*4*	*5*
Proportion of refixations	Silent	High	*M*	.360	.261	.196	.153	.157
			SE	.006	.006	.005	.005	.005
		Low	*M*	.398	.334	.262	.205	.213
			SE	.006	.006	.005	.005	.005
	Aloud	High	*M*	.353	.267	.191	.155	.165
			SE	.007	.006	.005	.005	.005
		Low	*M*	.382	.288	.224	.182	.190
			SE	.007	.006	.005	.006	.005
Proportion of regressions	Silent	High	*M*	.138	.127	.117	.107	.112
			SE	.006	.005	.005	.005	.005
		Low	*M*	.166	.163	.135	.136	.152
			SE	.006	.006	.005	.006	.005
	Aloud	High	*M*	.152	.137	.120	.115	.121
			SE	.007	.006	.005	.006	.006
		Low	*M*	.152	.133	.119	.114	.119
			SE	.007	.006	.005	.005	.006

materializes starting in second grade and is then maintained through higher grades (Figure 2c).

Proportion of regressions.   Again, we found a significant main effect for Word Frequency ($b = 0.014$, $SE = 0.002$, $t = 6.57$) with more regressions on low frequency words. This main effect was qualified by a significant Reading Mode by Word Frequency interaction ($b = -0.032$, $SE = 0.004$, $t = -7.29$). Word Frequency affected regressions only in silent reading but not when reading aloud (see Figure 2d). No other effects were significant (all ts < 1.5).

Effects of comprehension.   For additional analysis, comprehension level was calculated for each participant, based on the answers to comprehension questions during the task. Due to technical issues, comprehension scores were only available for 464 participants. This sample was divided into stronger and weaker comprehenders (median-split by grade) and Comprehension Level was then added as a 2-level between subject factor into the model. While this procedure is not ideal, it allows gaining a first impression of comprehension level effects that will be elaborated in future analyses, when eye movement data will be combined with the psychometric assessment scores. In this paper we will restrict the analyses to total reading time and only report effects that were significantly affected by Comprehension Level.

TABLE 7
Means (*M*) and Standard Deviation (*SD*) for comprehension scores by Grade, Reading Mode, and Comprehension Level

			Grade				
			1	*2*	*3*	*4*	*5*
Silent	Stronger	*M*	11.40	15.81	16.38	18.61	19.70
		SD	3.71	2.90	3.01	1.65	1.47
	Weaker	*M*	2.24	6.46	10.45	13.83	12.95
		SD	1.77	3.47	3.05	3.20	4.37
Aloud	Stronger	*M*	14.40	16.47	17.25	17.70	18.27
		SD	2.70	2.03	1.87	1.40	1.21
	Weaker	*M*	5.65	10.42	12.56	13.88	13.02
		SD	3.21	4.09	2.42	1.51	3.28

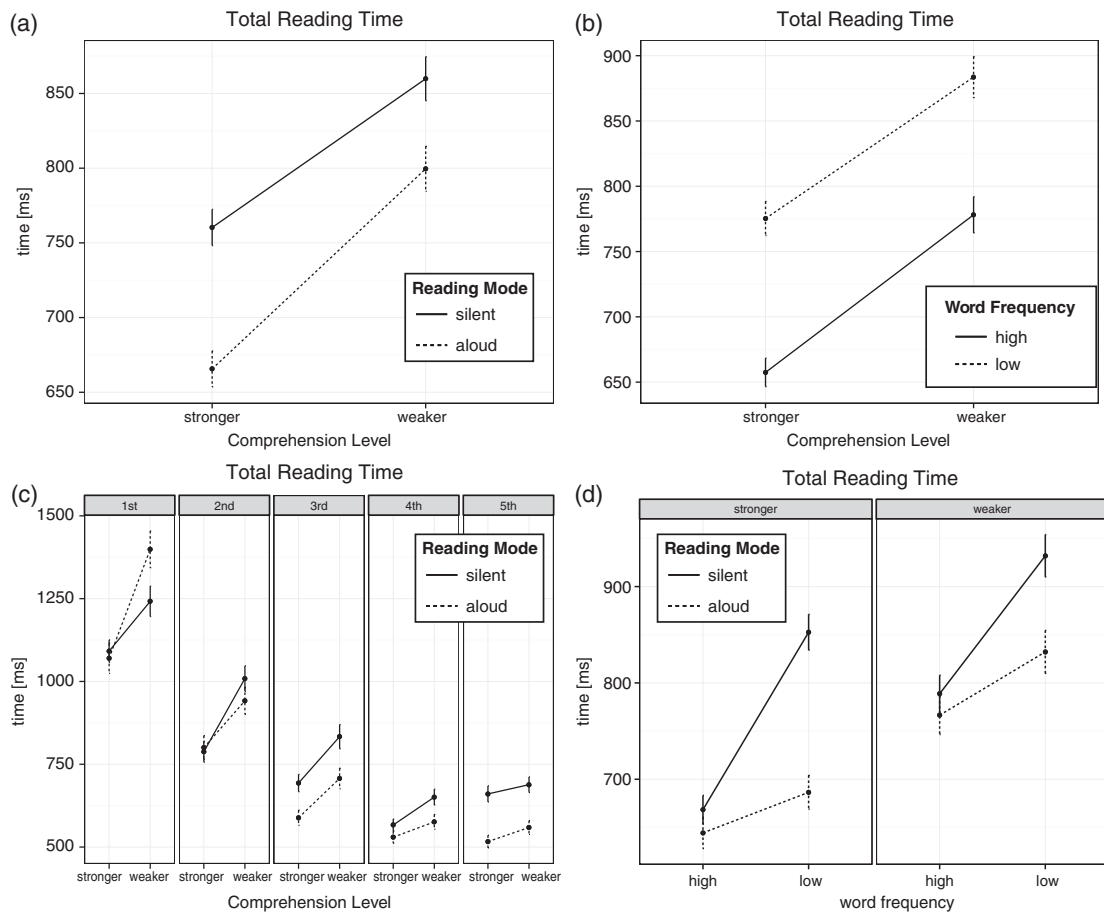Maximum possible score = 24 points.

Figure 3. Effects of differences in comprehension level on total reading time in ms (*M* and *SE*). (a) Comprehension Level by Reading Mode interaction, (b) Comprehension Level by Word Frequency Interaction, (c) Grade by Comprehension Level by Reading Mode interaction, and (d) Comprehension by Reading Mode by Word Frequency interaction.

Mean comprehension for weaker and stronger comprehenders are reported by grade in Table 7. Although there was no significant main effect of Comprehension Level, there was a significant interaction between Comprehension Level and Reading Mode ($b = 0.060$, $SE = 0.015$, $t = 3.97$). Reading aloud led to a larger increase in reading times for weaker comprehenders compared to stronger comprehenders (Figure 3a). The Comprehension Level by Word Frequency interaction was also significant ($b = -0.024$, $SE = 0.012$, $t = -2.00$), with a frequency effect that was slightly less pronounced for weaker comprehenders (Figure 3b).

A significant Grade x Comprehension Level x Reading Mode interaction between 1st and 2nd grade ($b = -0.254$, $SE = 0.052$, $t = -4.90$) results from the fact that weaker comprehenders in 1st grade had prolonged reading times when reading aloud, whereas stronger comprehenders showed a small reduction (see Figure 3c). From 2nd grade on, both stronger and weaker comprehenders had shorter times when reading aloud, although weaker comprehenders always needed additional time.

Finally, the significant Comprehension Level x Reading Mode x Frequency interaction ($b = 0.074$, $SE = 0.024$, $t = 3.07$) indicated that the influence of reading mode on the word frequency effect was greater for stronger compared to weaker comprehenders and even more pronounced for low frequency words (see Figure 3d).

DISCUSSION

The number of publications on eye movement behaviour in children has increased substantially over the last decade, especially with regard to experimental studies of word processing, and, to a lesser extent, higher-level processing on the sentence and text level (Blythe & Joseph, 2011). However, it is evident that we are still far away from a comprehensive understanding of reading development as expressed in the dynamics of eye movement behaviour. This paper reports the first results of a large-scale combined cross-sectional and longitudinal project that will eventually integrate oculomotor data with results from a large battery of psychometric assessments (Kuperman & Van Dyke, 2011).

The present work used identical sentence materials for grades 1 to 5. This approach was taken to avoid the complexities of age-appropriate or grade-appropriate materials, where it is difficult to dissociate changes due to development from differences induced by making text more complex (see Blythe & Joseph, 2011, and Radach et al., 2008, for discussions). It could be argued that even sentences that are carefully designed with an eye on grade-adjusted frequency norms (Zeno et al., 1995) must necessarily become increasingly less difficult to read, creating a qualitatively different task setting. Our position in this debate is that even though it is problematic to compare word frequency contrasts

between identical words across grades, this comparison comes with the advantage that the incoming visual and linguistic information is virtually identical. Results therefore reflect an upper limit to developmental effects, without being affected by any additional source of variance.

General developmental trends

The two basic oculomotor parameters determining reading times are the duration of fixations and the number of fixations. There is no doubt that fixation durations reflect the mental workload associated with linguistic processing in reading, even though opinions may differ on the extent of cognitive control (O'Regan, Vitu, Radach, & Kerr, 1994; Rayner, Sereno, & Raney, 1996). Blanchard (1985) was the first to show that in adult data increases in word processing difficulty are even more strongly related to the number of fixations, leading to the convention of using gaze duration as the primary measure in the vast majority of published studies (Rayner, 1998). We are reporting word viewing durations in terms of initial fixation durations, refixation time and rereading time, allowing for a complete decomposition on the processing time line.

Looking at the general trajectory of development, our results confirm the expected pattern of faster reading with advancing development. Most notably, across all viewing time measure parameters the biggest gain appears to be happening between 1st and 2nd grade, whereas effects seem to level off between 4th and 5th grade[2]. Hand in hand with this decrease of viewing durations goes the reduction of fixation probabilities and refixation frequencies. This is not surprising, as decoding skills increase over time, making additional fixations on a word less necessary in higher grades. At the same time, the proportion of interword regressions increases for higher grades. This finding might seem counterintuitive on first glance; however, it is in line with McConkie et al. (1991) and appears to reflect changes in reading strategy. More experienced readers do not have to spend as much time on words during first pass reading, but then selectively go back to words that might have caused processing difficulty. This is one way of making the reading process more efficient, a developmental tendency that is reflected in the observed changes of refixation and regression proportions (see also Vorstius et al., 2013).

We are in sympathy with the idea that the primary driving force of reading development is the increasing efficiency of lexical processing (Perfetti, 2007; Reichle et al., 2013). However, the large reduction of rereading time also points

---

[2] The levelling off between 4th and 5th grade is unlikely to be based on the sentences being too easy to read for most 5th grade students. As Table 7 indicates, the mean comprehension score of the weaker comprehension subgroup was only 13 of the possible 24 points. When analyses were restricted to this subsample of low performing children, developmental trends remained basically identical.

towards post-lexical processing on the sentence level as a major arena of developmental progress (see Huestegge et al., 2009, for similar data). It remains to be seen to what extent non-linguistic processing aspects of information processing in reading (attention, visuomotor control, spatial navigation, etc.) also contribute to the observed changes in eye movements. To allow for analyses of such influences, the present project also collected data in a reading-like scanning task (Hillen et al., 2013), which we will report in a later publication.

Differences between reading modes

The present work constitutes the first large-scale corpus-based comparison of eye movements in silent vs. oral reading (see Ashby, Yang, Evans, & Rayner, 2012; Inhoff & Radach, this volume, for comparisons of adult data in experimental designs). Initial fixation durations and refixation duration were longer in oral reading, but decreased across grades in both reading modes. A decrease was also found for rereading times, however, this reduction was more pronounced when reading aloud, especially between 1st and 2nd as well as 2nd and 3rd grade. Interestingly, this resulted in rereading times that were actually higher in silent reading for higher grades. We will return to this fascinating finding in the discussion below on effects of word frequency.

A further decomposition of fixation durations into single fixation duration, first of two, and second of two fixation durations yielded a very interesting result. Irrespective of whether a fixation was the first or second in a sequence of two, and irrespective of whether this sequence was part of the initial gaze (first pass reading) or later rereading, a very similar increase in fixation duration was found in oral as compared to silent reading. This means that no matter what the functional role of a fixation is over the time course of word processing (from early orthographic processing up to lexical access and re-analysis), there is a comparable delay in the initiation of the following saccade. In our view the most likely interpretation of this pattern is an overall slowing in the speed of processing due to increased capacity demands from concurrently active processing modules. Candidates for mechanisms causing this bottleneck include speech planning, articulation monitoring and eye–voice coordination, which may all impose extra demands on working memory during oral reading (Inhoff, Connine, Eiter, Radach, & Heller, 2004; Inhoff et al., 2011).

Looking at effects related to the number of fixations and their distribution, refixation rates were higher, whereas the proportion of interword regressions was lower when reading aloud. Most likely the higher frequency of refixations reflect an adaptive adjustment, since reading aloud requires the reader to articulate every word, which may lead to more emphasis on sublexical word properties related to phonological and articulation-related word characteristics (e.g., Ashby, Treiman, Kessler & Rayner, 2006; Breen & Clifton, 2011; Huestegge, 2010). The substantially lower proportion of interword regression in oral reading is

likely to be a consequence of constraints imposed by eye–voice coordination. Even though the generation of regressive saccades is one way of dealing with instances when the eyes get too far away from the currently articulated word, it appears that the need to maintain an optimal eye–voice span eliminates regressive saccades that would normally be made in silent reading (Inhoff et al., 2011). This raises the interesting question of whether the suppression of regressions might be detrimental to comprehension in reading aloud.

Effects of word frequency

Starting with initial fixation durations, overall target word analyses closely mirrored the findings obtained from the corpus of all words. Initial fixation duration decreased across grades, and fixation durations were also affected by word frequency, leading to longer fixation durations on low frequency words. For refixation times on target words, there was again a decrease across grades.

Interestingly, however, the reading mode effect for refixation times was reversed for target words. Compared to those over all words, less time was spent refixating targets in the oral reading mode (see Figure 2a). This effect was even more pronounced for rereading times. Especially in higher grades, rereading times are substantially longer in silent reading, which is consistent with the respective results from the full word sample. Taken together, these effects result in more total viewing time spent on target words when reading silently. Assuming that even high frequency target words (usually agents or objects of an event or action) are important relative to the many other words in the rest of the sentence, this result appears to indicate that during oral reading word viewing may not be focused as much on words that are most critical for comprehension. More generally, then, viewing durations during reading aloud might be less sensitive to moment to moment variation in the cognitive effort related to word recognition. This conclusion is in line with the reading mode differences in word frequency effects discussed below.

Word frequency effects on viewing time measures tended to be attenuated in reading aloud, even though the respective main effect did not reach significance. However, the significant three-way Grade by Reading Mode by Frequency interaction for refixation times indicates that the effect of Reading Mode is more pronounced for low compared to high frequency words in 2nd vs. 1st grade. This effect is based on an attenuated influence of word frequency when reading aloud in second (and subsequent) grades. As suggested above, in oral reading the decision to terminate a fixation and initiate the next saccade may be co-determined by word properties related to pronunciation and the coordination of word processing and speech production is likely to be another significant source of variability, especially in the lowest grade.

In addition, the proportion of refixations showed a matching pattern. Fewer refixations were made on high frequency words and this effect was also modulated

by a Reading Mode by Word Frequency interaction. The frequency effect was more pronounced when reading silently. Also matching the temporal data was the three-way interaction between Grade, Reading Mode and Word Frequency, showing that dissociation between reading modes for low frequency words starts in 2nd grade and is maintained thereafter. Furthermore, the proportion of regression also showed a reliable Reading Mode by Word Frequency interaction with no effect of word frequency when reading aloud but a higher proportion of regressions to low frequency words in silent reading. Apparently the processing of low frequency words leads to a much larger number of regressive saccades, in line with the idea that difficult words will more often require re-analysis. The virtual absence of this effect in oral reading is stunning and supports the hypothesis that constraints of eye–voice coordination may indeed suppress regressions back to difficult words.

Above and beyond these general effects, a very interesting pattern of results emerged for rereading times (see Table 5). Evidently, oral rereading times in 1st grade were inflated, independent of word frequency. This effect transitions in 2nd grade and is reversed for grade 3 and higher. We suggest that this qualitative difference is related to substantial difficulties experienced by many beginning readers with the language production component of oral reading. It is quite likely that many of these young readers still have articulation difficulties and their routines for grapheme–phoneme conversion are not yet automatized enough to allow for smooth processing within the given time line constraints.

Effects of reading comprehension

Eye movement analyses have proven useful in prior research to illuminate comprehension processes as they unfold over time during reading (e.g., Rayner, Chace, Slattery, & Ashby, 2006; Vorstius et al., 2013). In the present work we took a different perspective, asking how reading in the silent and oral mode differs as a function of comprehension as an individually acquired skill. Our analyses of relations between reading mode and comprehension focused on total reading time as an indicator of overall processing effort, comparing performance of more and less successful participants.

Weaker comprehenders, while reading substantially slower, showed a smaller word frequency effect. More specifically, for high frequency words the slowing of word reading times was greater than for low frequency words, when compared to stronger comprehenders. This effect could be the result of less frequent exposure to text, resulting in a less elaborate mental lexicon, hence not allowing faster decoding and integration into the text of high frequency words (see Hiebert & Fisher, 2007; Priya & Wagner, 2009, for discussions of vocabulary development).

With respect to Reading Mode, total reading times were longer for weaker comprehenders, but this effect was more pronounced when reading aloud, leading

to a decreased reading mode effect for weaker comprehenders. Furthermore, the interaction between Grade, Comprehension Level and Reading Mode indicated that this pattern was most pronounced for 1st grade, whereas higher grades showed no interaction pattern for Comprehension Level and Reading Mode (Figure 3c). After 3rd grade, total reading times in silent reading were longer in both groups. Here the reduced relative difficulty of reading materials may come into play, reducing the challenge posed by unfamiliar and irregular words.

This idea is in harmony with the effect of word frequency as evident in the Comprehension Level by Reading Mode by Word Frequency interaction. There was no difference between stronger and weaker comprehenders between reading modes for high frequency words. However, stronger comprehenders seem to allocate their resources more strategically when reading silently, i.e., spending more time on difficult low frequency words. This results in a stronger increase in total viewing time on low frequency words when reading silently for stronger comprehenders, while the frequency effect was reduced when reading aloud. Although the frequency effect when reading aloud was also attenuated in weaker comprehenders, the increase was slightly steeper (compared to stronger comprehenders), whereas the increase in the silent reading mode was somewhat less pronounced. Importantly, differences between reading modes were virtually identical on high frequency words for both stronger and weaker comprehenders. The steeper increase in weaker comprehenders when reading aloud could be taken to indicate that the articulation of less familiar words affects children with less developed comprehension skills more strongly.

Taken together, it appears that these data reflect the fact that weaker comprehenders have not yet acquired the flexibility needed to take advantage of the more unconstrained nature of silent reading, allowing allocation of processing time to more difficult words when needed. In other words, comprehension difficulties may be more likely when silent reading functions more like an internalized version of reading aloud. This tentative conclusion sets an interesting counterpoint to the suggestion by Juel and Holmes (1981) that oral reading may have differential effects on comprehension depending on individual progress in reading development. Both ideas appear well suited to be addressed directly in well-controlled experimental research.

REFERENCES

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 416–424.

Ashby, J., Yang, J., Evans, K. H. C., & Rayner, K. (2012). Eye movements and the perceptual span in silent and oral reading. *Attention, Perception, & Psychophysics*, *74*, 634–640.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*, 390–412.

Baayen, R. H., Piepenbrock, R., & van Rijn H. (1993). *The CELEX lexical database (Release 1)* [CD-ROM]. University of Pennsylvania, Philadelphia, PA: Linguistic Data Consortium.

Bates, D., Mächler, M., & Bolker, B. (2012): lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0.

Blanchard, H. E. (1985). A comparison of some processing measures based on eye movements. *Acta Psychologica*, *58*, 1–15.

Blythe, H. L., Häkiö, T., Bertram, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research*, *51*, 84–92.

Blythe, H. L., & Joseph, H. S. S. L. (2011). Children's eye movements during reading. In S. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 643–662), Oxford: Oxford University Press.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular coordination of eye movements during reading in children and adults. *Vision Research*, *46*, 3898–3908.

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.

Breen, M., & Clifton, C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language*, *64*, 153–170.

Buswell, G. T. (1922). *Fundametal reading habits, a study of their development*. Chicago, IL: Chicago University Press.

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dymamical model of saccade generation during reading. *Psychological Review*, *112*, 777–8I3.

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.

Hiebert, E. H., & Fisher, C. W. (2007). The critical word factor in texts for beginning readers. *Journal of Educational Research*, *101*(1), 3–11.

Hillen, R., Günther, T., Kohlen, C., Eckers, C., van Ermingen-Marbach, M., Sass, K. Scharke, W., Vollmar, J., Radach, R., & Heim, S. (2013, epub ahead of print). Identifying brain systems for gaze orienting during reading: fMRI investigation of the Landolt paradigm. *Frontiers in Human Neuroscience*.

Huestegge, L. (2010). Effects of vowel length on gaze durations in silent and oral reading. *Journal of Eye Movement Research*, *3*, 1–18.

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research*, *49*, 2948–2959.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology: Learning Memory and Cognition*, *21*, 1430–1440.

Inhoff, A. W., Connine, C., Eiter, B., Radach, R., & Heller, D. (2004). Phonological representation of words in working memory during sentence reading. *Psychonomic Bulletin & Review 11*, 320–325.

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 29–54). Oxford: Elsevier.

Inhoff, A. W., Solomon, M., Radach, R., & Seymour, B. (2011). Temporal dynamics of the eye voice span and eye movement control during oral reading. *Journal of Cognitive Psychology*, *23*, 543–558.

Inhoff, A. W., & Weger, U. W. (2003). Advancing the methodological middle ground. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movements*. Oxford: Elsevier Science.

Inhoff, A. W., Weger, U. W., & Radach, R. (2005). Sources of information for the programming of short- and long-range regressions during reading. In Underwood, G. (Ed.), *Cognitive Processes in Eye Guidance*. Oxford: Oxford University Press.

Jones, E. E., & Lockhart, A. V. (1919). A study of oral and silent reading in the elementary schools of Evanston. *School and Society*, *10*, 587–590.

Joseph, H. S. S. L., & Liversedge, S. P. (2013). Children's and adults' processing of syntactically ambiguous sentences during reading. *PLoS ONE*, *8*(1): e54141.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., Gathercole, S. E., & Rayner, K. (2008). Children's and adults' processing of anomaly and implausibility during reading: Evidence from eye movements. *Quarterly Journal of Experimental Psychology 61*, 708–723.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, *49*, 2078–2086.

Joseph, H. S. S. L., 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.

Juel, C., & Holmes, B. (1981). Oral and silent reading of sentences. *Reading Research Quarterly*, *16*(4), 545–568.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*, 262–284.

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, *18*, 655–681.

Kragler, S. (1995). The transition from oral to silent reading. *Reading Psychology: An International Quarterly*, *16*, 395–408.

Krügel, A., & Engbert, R. (2010) On the launch-site effect for skipped words during reading. *Vision Research*, *50*, 1532–1539.

Kuperman, V., & Van Dyke, J. A. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, *65*, 42–73.

McConkie, G. W. (1981). Evaluating and reporting data quality in eye movement research. *Behavior Research Methods & Instrumentation*, *13*, 97–106.

McConkie, G. W., & Dyre, B. P. (2000). Eye fixation durations in reading: Models of frequency distributions. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 683–700). Oxford: Elsevier.

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

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception and Psychophysics*, *46*, 245–253.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movement during reading. In J. F. Stein (Ed.), *Vision and visual dyslexia* (pp. 251–262). London: Macmillan Press.

Miller, S. D., & Smith, D. E. P. (1990). Relations among oral reading, silent reading and listening comprehension of students at different competency levels. *Reading Research and Instruction*, *29*, 73–84.

O'Regan, J. K., Vitu, F., Radach, R., & Kerr, P. (1994). Effects of local processing and oculomotor factors in eye guidance in reading. In Ygge, J., & G. Lennerstrand (Ed.), *Eye movements in reading*. New York, NY: Pergamon.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383.

Priya, K., & Wagner, R. K. (2009). The roles of fluent decoding and vocabulary in the development of reading comprehension. In R. K. Wagner, C. L. Schatschneider, & C. Phythian-Sense (Eds.),

Beyond decoding: The behavioral and biological foundations of reading comprehension (pp. 124–139). New York: Guilford Press.

Prior, S. M., & Welling, K. A. (2001). "Read in your head": A Vygotskian analysis of the transition from oral to silent reading. *Reading Psychology*, 22(1), 1–15.

Pynte, J., & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. *Vision Research*, 46(22), 3786–3801.

R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye movement control during reading. *Psychological Research*, 72, 675–688.

Radach, R., & Kempe, V. (1993). An individual analysis of initial fixation positions in reading. In d'Ydevalle, G., & van Rensbergen, J. (Eds.), *Perception and cognition. Advances in eye movement research*. Amsterdam: Elsevier North Holland.

Radach, R., & Kennedy, A. (2004). Theoretical perspectives on eye movements in reading: Past controversies, current issues, and an agenda for the future. *European Journal of Cognitive Psychology*, 16, 3–26.

Radach, R., & Kennedy, A. (2013). Eye movements in reading: Some theoretical context. *Quarterly Journal of Experimental Psychology*, 66, 429–452.

Radach, R., & McConkie, G. W. (1998). Determinants of fixation positions in words during reading. In G. Underwood (ed.), *Eye guidance in reading and scene perception*. Oxford: Elsevier.

Radach, R., Schmitten,Chr.,Glover, L., & Huestegge, L. (2008). How children read for comprehension: Eye movements in developing readers In R. K. Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension*. New York: Guilford Press.

Rayner, K. (1979). Eye guidance in reading: fixation locations within words. Perception, 8, 21–30.

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. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8), 1457–1506.

Rayner, K., Chace, K. H., Slattery, T. J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, 10, 241–255.

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

Reichle, E. D., Liversedge, 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, 110–149.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, 105, 125–157.

Reilly, R., & Radach, R. (2006). Some empirical tests of an interactive activation model of eye movement control in reading. *Cognitive Systems Research*, 7, 34–55.

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, 1270–1281.

Swalm, J. E. (1973). A comparison of oral reading, silent reading, and listening comprehension. *Education*, 92, 111–115.

Tang, S., Reilly, R. G., & Vorstius, C. (2012). EyeMap: A software system for visualizing and analyzing eye movement data in reading. *Behavioral Research Methods*, 44, 420–438.

Taylor, S. E., Frackenpohl, H., & Pettee, J. L. (1960). Grade level norms for the components of fundamental reading skill. In *EDL Research and Information Bulletin* (Vol. 3). Huntington, NY: Educational Development Laboratories.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. (4th Edn.). Berlin: Springer.

Vitu, F., & McConkie, G. W. (2000). Regressive saccades and word perception in adult reading. In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 301–326). Oxford: Elsevier.

Vorstius, C., Radach, R., Mayer, M., & Lonigan, C. (2013). Monitoring local comprehension monitoring in sentence reading. *School Psychology Review*, *42*, 191–206.

Yang, S-N., & McConkie, G. W. (2001). Eye movements during reading: A theory of saccade initiation times. *Vision Research*, *41*, 3567–3585.

Zeno, S., Ivens, S., Millard, R., & Duvvuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science Associates.