Children's Eye-Movements During Reading Reflect the Quality of Lexical Representations: An Individual Differences Approach

Steven G. Luke, John M. Henderson, and Fernanda Ferreira

Children's Eye-Movements During Reading Reflect the Quality of Lexical Representations: An Individual Differences Approach

Steven G. Luke
Brigham Young University

John M. Henderson and Fernanda Ferreira
University of California, Davis

The lexical quality hypothesis (Perfetti & Hart, 2002) suggests that skilled reading requires high-quality lexical representations. In children, these representations are still developing, and it has been suggested that this development leads to more adult-like eye-movement behavior during the reading of connected text. To test this idea, a set of young adolescents (aged 11–13 years) completed a standardized measure of lexical quality and then participated in 3 eye-movement tasks: reading, scene search, and pseudoreading. The richness of participants' lexical representations predicted a variety of eye-movement behaviors in reading. Further, the influence of lexical quality was domain specific: Fixation durations in reading diverged from the other tasks as lexical quality increased. These findings suggest that eye movements become increasingly tuned to written language processing as lexical representations become more accurate and detailed.

Keywords: reading, individual differences, reading development, eye movements

Skilled adult readers make eye movements, called *saccades,* about four times a second (Rayner, 1998, 2009). In between these movements are pauses called *fixations,* during which the eyes take in information about the text. Developing readers also make fixations and saccades, but the eye movements of child readers differ quantitatively from those made by adults: Children read at a slower rate, they make more and longer fixations (Blythe, 2014; Blythe & Joseph, 2011), they are more likely to look at words again before leaving them and to come back to them later (Blythe et al., 2011, 2009; Joseph et al., 2009), and they appear to gather information from a narrower window of the text (Rayner, 1986). This differences are observed even though children appear to be just as efficient as adults at extracting visual information from text (Blythe et al., 2011, 2009).

A potential explanation for these differences is that children's eye movements reflect their immature linguistic skills. In other words, oculomotor specialization for reading is driven by other important developmental changes related to increasingly sophisticated knowledge of language: As they age, children's vocabulary becomes larger, and meanings become more precise (Verhaeghen, 2003); children acquire mastery of more complex syntactic structures (Leech, Aydelott, Symons, Carnevale, & Dick, 2007; Montgomery, Magimairaj, & O'Malley, 2008; Vasilyeva, Waterfall, & Huttenlocher, 2008; Wassenberg et al., 2008); and they develop better metalinguistic awareness of linguistic conventions and style

techniques (Smith & Tager-Flusberg, 1982). These changes would be expected to influence how eye movements in reading change during early development: Oculomotor behaviors that are under cognitive control will become more responsive to linguistic factors (and more adult like) as language skill increases. Skilled reading clearly recruits the language system, as indicated by strong influences of word- and sentence-level variables on eye movements in reading (Just & Carpenter, 1980; Kliegl, Nuthmann, & Engbert, 2006; Rayner, 1998), and individual differences in verbal skills influence eye-movement behavior even in adults (Kuperman & van Dyke, 2011), so it seems reasonable to suppose that the development of the language system might lead to increasingly skilled reading.

Although language develops in a variety of ways throughout childhood and adolescence, there are reasons to believe that vocabulary-related skills are of primary importance in the maturation of reading skill. Rate of lexical processing appears to be the largest determinant of differences between child and adult reading (Reichle et al., 2013). This observation is consistent with the Lexical Quality Hypothesis (Hamilton, Freed, & Long, 2013; Perfetti, 2007; Perfetti & Hart, 2002), which assumes that successful reading is based on the availability of high-quality lexical representations, which tightly link orthographic, phonological, and semantic word details.

The idea that increased linguistic proficiency drives changes in eye-movement behavior, though intuitive and appealing, is not the only possible explanation for differences in reading skill between children and adults. It is also possible that developmental changes in the oculomotor system itself allow readers more precise control of eye movements. Reading is, after all, a visual task as well as a linguistic one, and, in adults, aggregate eye-movement patterns in reading appear to correlate with those in other tasks (Henderson & Luke, 2014; but cf. Rayner, Li, Williams, Cave, & Well, 2007). Additionally, some of what happens during reading is largely independent of cognitive processes (Luke & Henderson, 2013;

Steven G. Luke, Department of Psychology, Brigham Young University; John M. Henderson, Department of Psychology and Center for Mind and Brain, University of California, Davis; Fernanda Ferreira, Department of Psychology, University of California, Davis.

Correspondence concerning this article should be addressed to Steven G. Luke, Department of Psychology, Brigham Young University, 1001 Spencer W. Kimball Tower, Provo, UT 84602. E-mail: steven_luke@byu.edu

Nuthmann, Engbert, & Kliegl, 2007; Rayner & Fischer, 1996; Vitu, O'Regan, Inhoff, & Topolski, 1995), and the eyes respond to processing difficulty similarly in reading and nonreading tasks (Luke, Nuthmann, & Henderson, 2013). Saccade control, as measured by saccade latency and accuracy, continues to improve throughout childhood and into adolescence (Fukushima, Hatta, & Fukushima, 2000; Salman et al., 2006), as does cognitive control of eye movements (Luna, Velanova, & Geier, 2008). Such changes would be expected to be beneficial in reading, which requires precise and rapid control of eye movements and therefore could explain differences between child and adult readers.

It is of course possible (and indeed, likely) that both of these accounts are correct: The transformation from novice to skilled reader may require both linguistic and oculomotor development. To fully understand how children develop into proficient readers, these two types of changes must be carefully dissociated, so that both can be studied and understood (Blythe & Joseph, 2011). This means that any attempt to evaluate the lexical quality hypothesis must control for potential differences in oculomotor development. Fortunately, linguistic and oculomotor development should influence eye movements in different ways. Specifically, any influence of individual differences in linguistic skill should be task specific, influencing eye movements in reading but not in other visual tasks in which language has little or no involvement. The development of oculomotor skill, on the other hand, should be more task general, leading to more efficient eye movements in most visual tasks.

The goal of this study was to investigate the influence of linguistic skill, specifically the quality of children's lexical representations (as measured by scores on a standardized reading comprehension test) on eye-movement behaviors in reading. If individual differences in lexical quality can account at least in part for differences in reading eye movements, then participants with more precise lexical representations should be more efficient readers, with shorter fixations and/or changes in saccade targeting (more skipping, fewer word refixations, fewer regressions). As mentioned above, it is important to dissociate the potential influences of linguistic and oculomotor skill when studying reading development. We controlled for the influence of oculomotor development in two ways. First, we tightly controlled the ages of the participants: The average age was 11.8 years old, with a range of 11–13 ($SD = 0.58$). By about 11 years old, children are beginning to be mostly adult like in their reading behaviors (Blythe, 2014; Blythe & Joseph, 2011). Further, there is evidence that saccade control is mostly mature in 11-year-old children (Fukushima, Hatta, & Fukushima, 2000), whereas in younger children there is a greater likelihood of variability in global oculomotor development. Thus, by selecting early adolescents in a narrow age range, any differences in reading skill related to maturation of oculomotor control should be minimal. Participants were also screened so that all were adequate word decoders; it is well established that children with dyslexia or other reading disorders differ significantly from other children in their eye movements during reading (Kirkby et al., 2011, 2008), and the goal of this study was to investigate normal reading development. Second, we compared the eye-movement behaviors of our participants in three different visual tasks: reading, pseudoreading, and visual search. If lexical quality does indeed underlie reading development, then adolescents with richer lexical representations should show more efficient (and adult-like)

reading behaviors. Further, any influences of lexical quality should be domain specific, not extending to other nonreading visual tasks.

Method

Participants

Twenty-one adolescents aged between 11 and 13 years (mean age 11.8 years) from Columbia, South Carolina, and the surrounding regions took part in the study. Children who participated completed a written assent form, and their parents completed a written consent form. All participants included in the study reported typical language development, with no history of a speech, language, or hearing disorder. All participants had normal or corrected-to-normal vision.

Individual Difference Measures

The following individual-difference measures were completed by the participants.

Word-Reading Efficiency. Participants completed the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), which is a test of oral word reading ability. The TOWRE requires that the child read aloud as quickly and accurately as possible a list of real and nonwords, which are ranked according to difficulty. The score is a combined measure of fluency and accuracy of decontextualized word reading. This measure was included to establish children's word decoding ability and to provide related exclusionary criteria: To ensure that participants were normal readers, participants had to score ≥ 1 SD of the mean for their grade level. Because of this exclusionary criterion, the mean of this group was 111, significantly higher than (but within 1 SD of) the TOWRE norms mean of 100, $t(20) = 3.35$, $p < .0032$. The standard deviation of scores in our sample was 14.8, similar to the TOWRE norm standard deviation of 15. Thus, although as a group these participants were somewhat better-than-average decoders, the variability in scores was comparable to test norms.

Lexical Richness. Participants completed two subtests that comprise the Reading Comprehension portion of the Woodcock Reading Mastery Tests, Third Edition (WRMT-III; Woodcock, 2011). In the Word Comprehension subtest, the examinee is asked to provide antonyms and synonyms for visually presented words as well as complete a set of verbal analogies (e.g., *Compose* is to *music* as *choreograph* is to . . .). Participants also completed the Passage Comprehension subtest of the WRMT-III. This test is designed to measure overall text comprehension. The examinee reads passages silently and is asked to fill in a missing word. The mean score on the Word Comprehension subtest was 49.24 (out of 86; $SD = 11.4$), and the mean score on the Passage Comprehension subtest was 27.67 (out of 38; $SD = 5.05$). Scores on these two subtests were combined to form the Lexical Richness score. Proportions correct on the two subtests were summed, and then z scores were computed.

These tasks require participants to produce a particular word from semantic cues, either contextual (passage comprehension) or lexical (analogy, synonym–antonym). Obtaining high scores therefore requires that lexical representations be semantically rich; if a participant's lexical representations are semantically underspecified, then semantic cues will be less helpful and performance on

these tasks will suffer. Further, lexical representations must be semantically and orthographically precise, to choose exactly the right word and not a similar one, and also well-integrated, so that participants can easily access the appropriate orthography in response to the available semantic cues. Thus, these tasks measure aspects of lexical quality that are highly consistent with the definition for quality given by Perfetti (2007, Table 1). This score was therefore chosen as a measure of the semantic richness of participants' vocabulary.

The participants' scores on the Word-Reading Efficiency and Lexical Richness measures were correlated to examine the relation between them. This correlation ($r = .25$) was not significant ($p = .27$).

Apparatus

Eye movements were recorded through an SR Research Eyelink 1000 eye tracker (spatial resolution 0.01°) sampling at 1000 Hz. Participants sat 90 cm away from a 20-in. monitor, so that scenes and texts subtended approximately 20° × 15° of visual angle. Head movements were minimized with chin and head rests. Although viewing was binocular, eye movements were recorded from the right eye. The experiment was controlled with SR Research Experiment Builder software. Eyetracking began with a 9-point calibration routine used to map eye position to screen coordinates. Eyetracker calibration was not accepted until the average error was less than .49° and the maximum error was less than .99°. Participants were recalibrated at the start of each block and as needed during testing.

Stimuli

Stimuli consisted of 48 photographs of visual scenes and 80 age-appropriate texts (see Figure 1).

Scenes. Scenes were full color photographs depicting indoor and outdoor real-world environments (Figure 1A). Two thirds of these scenes contained a letter embedded in them (12-point Tahoma font). The target letter was pseudorandomly embedded in each scene so that across scenes the letters were approximately spatially randomly distributed (Brockmole, Castelhano, & Henderson, 2006). For the remaining one third of the scenes, no target letter was present.

Texts and pseudotexts. The texts were taken from an online news magazine for adolescents and ranged from 40 to 60 words in length. Half of the texts were presented in a regular font (Figure 1B), whereas for the other half each letter in the text was replaced with a shape using a custom pseudofont that removed all meaning but retained the overall visual structure of the text (Henderson & Luke, 2012, 2014; Luke & Henderson, 2013). Figure 1C presents an example passage in the pseudofont. For the text and pseudotext, regions of interest were defined around each word or pseudoword so that standard reading measures could be derived.

Procedure

In each session, all participants completed the same three tasks in the following order: text-reading, scene search, and pseudoreading. The purpose of inserting the scene search task between text- and pseudoreading was to minimize transfer effects across the two

Figure 1. Example stimuli from the three different tasks. A: An example scene with search target circled. Search target was not circled during presentation. B: An example text from the reading task. C: An example pseudotext. See the online article for the color version of this figure.

reading tasks. The order of stimulus presentation within each task was identical for each participant. See Swets, Desmet, Hambrick, and Ferreira (2007) for an account of why task and trial order is kept constant, rather than randomized or counterbalanced, across participants in individual difference studies. In the text-reading task, participants were instructed to read the 40 passages at their normal reading pace. In the scene search task, participants were instructed to search for an embedded target letter in 48 scenes. Participants were told that if they found the target, they should indicate this by fixating on the letter and pressing the response button. In pseudoreading, participants were instructed to move their eyes through the 40 pseudotexts "as if they were reading." These instructions have previously been used for pseudoreading

conditions (sometimes called "mindless reading") and in adults produce eye-movement behavior that approximates many of the eye-movement characteristics observed in reading (Henderson & Luke, 2012; Luke & Henderson, 2013; Nuthmann, Engbert, & Kliegl, 2007; Rayner & Fischer, 1996; Vitu, O'Regan, Inhoff, & Topolski, 1995).

Each trial in each task involved the following sequence. The trial began with a fixation marker at screen center for calibration check. Once a stable fixation was detected by the eyetracker, the stimulus was presented. The participant always viewed the stimulus for 12 s before it was removed from the screen, except in scene search trials when participants could end the trial earlier by pressing the response button (when they located the target letter). Following a trial, the screen was blanked and the next calibration screen appeared. If the calibration was not accurate, the participant was recalibrated and validated and then returned to the calibration screen.

Results

Individual Differences in Eye-Movement Behaviors in Reading

To establish that the individual difference measures did predict reading behavior, we performed a set of initial analyses on the reading data alone. These analyses investigated the relation between eye-movement behaviors during reading and reading skill as indexed by two individual differences measures: Word-Reading Efficiency (as measured by the TOWRE) and Lexical Richness (as measured by scores on the reading comprehension subtests of the WRMT-III; see Language Measures above). Age (in years and months) was also included as a predictor, to see if the minimal variation in age within our group was enough to influence reading behaviors.

Several different reading behaviors were investigated. First were measures of reading times. Three reading-time measures were analyzed: *first fixation duration, gaze duration,* and *total reading time.* First fixation duration is the duration of the first fixation on a word, whereas gaze duration is the sum of the durations of all fixations made on a word before leaving it the first time. Total time is the sum of the durations of all fixations on a word, including rereading. We hypothesized that participants with higher Lexical Richness scores would be more efficient readers and therefore would spend less time reading words according to one or more of these measures. In addition, we explored whether the well-established effects of word length, word frequency, and clause boundaries on reading times would be modulated by Lexical Richness. Given that these variables are often taken to influence lexical processing, we hypothesized that their influences would be modulated by Lexical Richness.

We also investigated three reading measures related to saccade targeting: *word refixations, word skipping,* and *regressions.* These binary measures reflect the probability of different saccade behaviors. Many words are fixated only once, but for those fixated more than once, a refixation is defined as a second fixation on a word before leaving it. During reading, many words are never fixated, but are instead skipped. A regression is a saccade that returns the eyes to a word after having passed it to the right. We hypothesized that readers with higher Lexical Richness scores would have less need to refixate words and might be more likely to skip words and less likely to return to them, making them more adult-like (Blythe, 2014).

Reading time measures were log transformed, and all continuous variables were centered (i.e., transformed so that the mean was 0). For these analyses, we used linear (for reading time measures) or logit (for binary measures) mixed-effects models. Models were fitted in a stepwise fashion, and predictors or interactions that were not significant were removed from the model. All models had random by-participant and by-word intercepts and any random slopes that contributed to the model as assessed by likelihood ratio tests. Note that most of the predictors are between-subjects and/or between-words, making them Level 2 predictors, so that ultimately only one or two random slopes were considered for each model— most models included only a by-word slope for Lexical Richness. Descriptive summary statistics for means and standard deviations of all analyzed reading measures are presented in Table 1.

In the analyses of reading times, there were no effects of the individual difference measures in first fixation duration or total time (all $ts < 1.61$, all $ps > .1$). In the analysis of gaze duration, Lexical Richness scores were predictive ($b = -.28$, $SE = 0.13$, $t = -2.16$, $p = .031$; difference in gaze duration between highest and lowest scorer $= 57.3$ ms), indicating that adolescents with higher Lexical Richness scores had shorter gaze durations. Neither Word-Reading Efficiency nor age was predictive of gaze duration (both $ts > 0.46$, both $ps > .64$).

The influence of word length on reading times was also investigated. In adult readers, word length has a strong influence on fixation times (Rayner, 2009). The goals of these analyses were, first, to see whether length effects were present in our adolescent participants and, second, to see whether the magnitude of these effects varied as a function of individual differences. For these analyses, words shorter than 4 and longer than 9 letters were eliminated, as they are either highly frequent or infrequent. Length was centered. There was a robust effect of word length on gaze duration and total time (both $ts > 3.64$, both $ps < .0003$), which did not interact with any individual difference measure (all $ts < 0.31$). In the analysis of first fixation duration, word length interacted with Lexical Richness score ($b = .27$, $SE = 0.0096$, $t = 2.79$, $p = .0053$), indicating that the effect of length was not

Table 1
Descriptive Statistics for the Different Reading Measures

Statistic	First fixation duration (ms)	Gaze duration (ms)	Total time (ms)	Refixation probability	Skipping probability	Regression probability
M	218	258	316	0.18	0.43	0.15
SD	101	166	231	0.38	0.5	0.36

significant for the low scorers but increased in size as Lexical Richness scores increased (see Figure 2).

An additional set of analyses was performed on the reading times to investigate the effects of word frequency. Word frequency has an influence on reading times in adults (Rayner, 2009) and children (Joseph, Nation, & Liversedge, 2013), so it is expected to be significant here. However, these is a possibility that the magnitude of this effect will vary as a function of individual differences in proficiency, as has been observed in second language learners (Diependaele, Lemhöfer, & Brysbaert, 2013). Frequencies for all words were derived from the COCA Corpus (Davies, 2009). For these analyses, all extremely short or long words (less than four, more than nine letters) and words for which no frequency was available were removed from the data. Word frequencies were regressed on word length, and the residuals from this regression were used to create residual frequency. This process, which removes the partial correlations between length and frequency, was performed because frequency and length are highly collinear: Long words tend to be less frequent, and the goal of the present analyses was to investigate independent effects of frequency. Residual frequency was a significant predictor of all reading time measures (all $ts > 4.47$, all $ps < .0001$), with shorter reading times for more frequent words. The effect of residual frequency did not interact with any individual difference measure (all $ts < 1.45$).

When readers encounter a clause or sentence boundary, they typically show inflated reading times (Just & Carpenter, 1980; Rayner, Kambe, & Duffy, 2000). This "wrap-up effect" is thought to reflect within- and between-clause integration costs. When the analyses for gaze duration and total time were rerun including punctuation (reflecting whether or not a word is at a clause or sentence boundary) as a predictor, wrap-up effects were observed in both measures (both $ts > 2.42$, both $ps < .016$; see Table 2). For gaze duration, the size of the wrap-up effect was not influenced by individual difference variables (all $ts < 1.14$). For total time, Lexical Richness (but not Word-Reading Efficiency or age; all

Table 2

Means (and SDs) Illustrating the Wrap-Up Effects in Gaze Duration and Total Time

Variable	Gaze duration (ms)	Total time (ms)
Words at clause or sentence boundaries	295 (227)	364 (324)
Other words	262 (163)	315 (225)
Wrap-up effect	33	49

$ts < 1.39$) modulated the effect of punctuation ($b = -.14$, $SE = 0.062$, $t = -2.2$, $p = .028$). Specifically, as Lexical Richness increased, the size of the wrap-up effect decreased, indicating that less skilled readers paused proportionally longer at clause and sentence boundaries than did more skilled readers (see Figure 3).

In sum, the analyses of reading times showed that readers with higher Lexical Richness scores had shorter gaze durations. Lexical Richness also modulated the early effect of word length and the clause wrap-up effect. We next report the results for the saccade-related eye-movement measures.

Refixation probability decreased significantly as Lexical Richness scores increased ($b = -1.01$, $SE = 0.39$, $z = -2.58$, $p = .01$). Neither Word-Reading Efficiency nor age influenced refixation probability (both $zs < 0.62$). This finding that more skilled readers were less likely to refixate a word is consistent with the overall lower gaze durations reported above (recall that gaze duration is the sum of the durations of all fixations on a word before leaving the word).

The analysis of skipping did not show any overall effect of any individual difference measure on skipping probability (all $zs < 0.46$, all $ps > .65$), indicating that more skilled readers showed no general tendency to skip words more or less often. An analysis that included word length as a predictor revealed that shorter words were more likely to be skipped ($b = -0.25$, $SE = 0.015$,
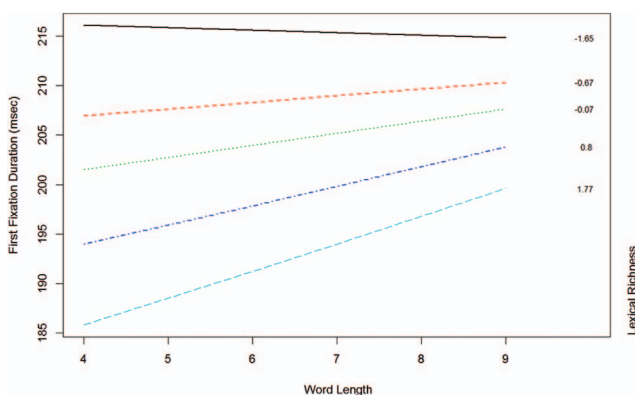
Figure 2. Interaction of word length and Lexical Richness score in the analysis of first fixation duration. Each line represents the influence of length on first fixation durations for a different quintile of Lexical Richness score (shown on right). Readers with lower quality lexical representations showed no length effect, with equally long first fixations for all word lengths, while readers with the highest-quality representations showed a significant effect of length, with shorter fixations on shorter words. See the online article for the color version of this figure.
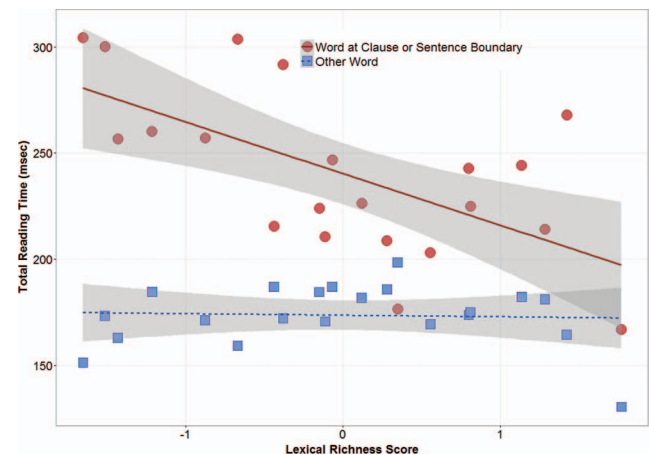
Figure 3. The interaction of the wrap-up effect with Lexical Richness score in the analysis of total reading time. The solid line represents total reading times at clause and sentence boundaries, and the dotted line represents total times at other words. Readers with higher quality lexical representations showed a smaller wrap-up effect. Error bands represent 95% confidence intervals. See the online article for the color version of this figure.

$z = -16.92$, $p < .0001$) and that this effect of length interacted with Lexical Richness score ($b = -0.025$, $SE = 0.0096$, $z = 2.59$, $p = .0097$). This interaction indicated that the effect of word length was weaker for adolescents with higher Lexical Richness scores; they were somewhat more likely to skip longer words than their peers with lower scores (see Figure 4). No other individual difference variables or interactions were significant (all zs < 1.04).

Regression probability was modulated by individual differences. Specifically, Lexical Richness scores were predictive of regressions, with more skilled readers more likely to make regressions ($b = .92$, $SE = 0.28$, $z = 3.27$, $p = .0011$). No effects or interactions involving Word-Reading Efficiency or age were significant (all ps > .13).

To summarize, individual differences in lexical richness, as measured by scores on the reading comprehension subtest of the WRMT-III, predicted a variety of eye-movement behaviors, including gaze durations, word refixations, and regressive saccades. Children are known to have longer gaze durations and to refixate words more often than do adults (Blythe et al., 2011, 2009; Joseph et al., 2009), therefore these findings indicate that adolescents with higher Lexical Richness scores have eye-movement patterns more similar to those of adults. Children with higher Lexical Richness scores also showed effects of word length on reading times earlier and showed a weaker effect of word length on skipping rates. Lexical Richness scores also predicted the size of clause wrap-up effects: the higher the score, the shorter the time spent at clause boundaries. Word-Reading Efficiency, a measure of oral word reading ability, did not influence eye-movement behavior in reading. Together these analyses suggest that eye movements in reading become more efficient with increased lexical richness. The only exception to this overall pattern arose in the analysis of regressions; adolescents with higher Lexical Richness scores made more regressions. We will return to this issue in the discussion.

Task Differences as a Function of Reading Skill

On the basis of the analyses conducted so far, it appears that eye movements in early adolescent readers are influenced by the
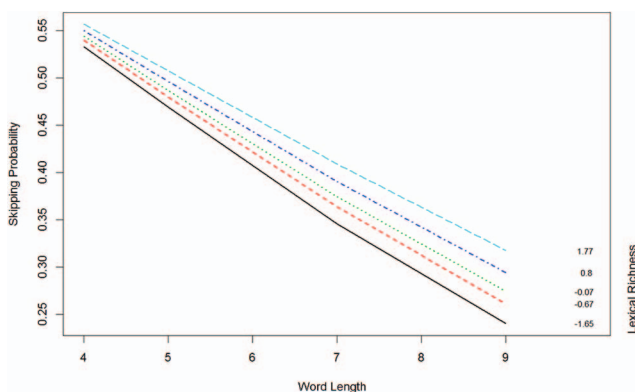

Figure 4. Interaction of word length and lexical richness score in the analysis of skipping probability. Each line represents the influence of length on skipping probability for a different quintile of lexical richness score (shown on right). Readers with lower quality lexical representations showed a stronger length effect, so that readers with the highest quality representations were more likely to skip longer words. See the online article for the color version of this figure.

quality of their lexical representations. There was no indication that orthographic decoding skill influenced reading behaviors in our group of proficient decoders. The goal of the next set of analyses, reported below, was to explore whether these individual differences are domain general or domain specific; in other words, we investigated whether higher scores on Lexical Richness (as measured by the two subtests of the WRMT-III) or on Word-Reading Efficiency (as measured by the TOWRE) reflect reading-specific skills or more general cognitive processes. To accomplish this, we analyzed differences in eye-movement behaviors among the three different tasks (reading, pseudoreading, and scene search) as a function of individual differences in reading skill. If a particular individual difference measure reflects a domain-specific skill, it should influence reading but not the other two tasks. On the other hand, if a measure reflects a skill that is not specific to reading, it should influence the other tasks instead of or in addition to reading.

The dependent variables in these analyses included fixation duration and saccade amplitude. The measures analyzed earlier, such as gaze duration or skipping rates, were not used here because such variables have no analogues in scene search that might permit a direct between-task comparison. Before analysis, fixations that contained blinks and their surrounding saccades were removed from the data, as were extremely short or long fixations (<50 ms or >1,500 ms). Extremely long saccades (>16 degrees) were also removed. Overall, about 6.4% of the data was discarded. Descriptive summary statistics for means and standard deviations of fixation duration and saccade amplitude in the three tasks are presented in Table 3.

For these analyses, we used linear mixed-effects models with individual fixation duration (log transformed) or saccade amplitude as dependent variables and task as a categorical predictor. Individual difference measures (Word-Reading Efficiency and Lexical Richness scores) were also included (as z scores), as was age, as continuous second-level predictors. All continuous predictors were centered. All models had random by-participant intercepts, and any random slopes that contributed to the model as assessed by likelihood ratio tests; all models included random by-participant slopes for task. Models were fitted in a stepwise fashion, and predictors or interactions that were not significant were removed from the model.

In the analyses of fixation durations, clear effects of task were observed, with significantly longer fixations in scene search and pseudoreading than in reading (both ts > 10.69, both ps < .001; see Table 3). This pattern is also observed in adults (Henderson & Luke, 2014; Luke, Nuthmann, & Henderson, 2013; Luke & Henderson, 2013; see also Kirkby et al., 2011). There were significant interactions between these task effects and Lexical Richness. These interactions indicate that the fixation duration difference between reading and the other two tasks increased as Lexical Richness scores increased (scene search, $t = 2.13$, $p = .033$; pseudoreading, $t = 2.22$, $p = .026$). These interactions are illustrated in Figure 5. No effects or interactions involving Word-Reading Efficiency or age were significant (all ts < 0.83).

While effects of task were observed in the analyses of saccade amplitude (see Table 3), with saccades in reading significantly shorter than in pseudoreading and scene search (both ts > 2.53, both ps < .012), no significant effects or interactions involving individual difference measures were observed (all ts < 1.57).

Table 3

Means (and SDs) for Fixation Duration (ms) and Saccade Amplitude (Degrees) for the Different Tasks

Variable	Reading	Pseudoreading	Scene search
Fixation duration	224 (104)	279 (166)	276 (128)
Saccade amplitude	3.15 (2.56)	3.53 (3.07)	3.84 (3.16)

In summary, adolescents had shorter fixations and saccades in reading than in the other tasks, consistent with previous research in adults (Henderson & Luke, 2014; Luke, Nuthmann, & Henderson, 2013; Luke & Henderson, 2013). For fixation durations, this difference was modulated by Lexical Richness, with more skilled readers showing a greater difference related to task.

Discussion

Children typically read more slowly than adults; they spend longer on each word, make more and longer fixations, and are more likely to look at a word a second time before moving on (Blythe, 2014; Blythe & Joseph, 2011). One potential explanation for the difference between adults' and children's eye-movement behavior during reading is that children's linguistic knowledge is still developing. Specifically, the Lexical Quality Hypothesis (Hamilton et al., 2013; Perfetti & Hart, 2002; Perfetti, 2007) proposes that successful reading requires high-quality lexical representations. In the present study, we assessed the quality of early adolescents' lexical representations by administering the reading comprehension subtest of the Woodcock Reading Mastery Tests—Third Edition. This subtest requires the participant to generate synonyms and antonyms and use context to supply missing words, making it a good proxy for lexical quality. These early adolescent participants then completed three active visual tasks, reading, pseudoreading, and scene search, while their eye movements were recorded.

Analyses of the participants' eye movements in reading revealed a significant influence of lexical richness on a variety of reading behaviors, such as gaze durations and word refixations. In these cases, children with more detailed lexical representations were more efficient in their eye movements. Furthermore, word length had an earlier effect on eye movements in participants with higher quality lexical representations, and these participants were also more likely to skip longer words. These findings provide evidence that lexical richness does indeed predict skilled reading.

When readers encounter a clause or sentence boundary, they typically show a "wrap-up" effect, an increase in reading times that is thought to reflect clause integration costs (Just & Carpenter, 1980; Rayner, Kambe, & Duffy, 2000). Another interesting finding from this study was that children with higher quality lexical representations showed a reduced wrap-up effect, suggesting a more efficient strategy of integrating information incrementally rather than at major clause boundaries. This effect was true for the analysis of total time, suggesting that more linguistically skilled children had less need to reread sentences and clauses. From a language-processing standpoint, this finding suggests that incremental processing is the optimal processing strategy and is associated with greater linguistic knowledge and skill; those who are not as good at comprehending need to pause more at punctuation boundaries and reread in order to integrate just-read content, but those who are more skilled do their processing on a more

word-by-word basis, incrementally integrating new material with what came before.

Two interesting exceptions in the analysis of the reading data should be noted. First, the influence of word frequency on reading times, though significant, was not modulated by lexical quality. It is important to note that the frequencies used here were obtained from an adult corpus, and therefore may be less representative of children's experience with words. When working with children and adolescents, it is important to carefully select both the items and the appropriate frequency measures before any definitive conclusions regarding frequency can be made (Joseph, Nation, & Liversedge, 2013). Also of interest was the effect of lexical quality on regressions, eye movements back to a previous word in the sentence. Children are generally more likely than adults to make regressions (Blythe, 2014). However, in the present data children with higher quality lexical representations made more regressions than did children with lower quality lexical representations, meaning that the more skilled readers appeared less adult like. Adults typically make regressions when they have comprehension difficulty (Rayner, 2009). One possible explanation for this discrepancy, then, is that the children in the present study are still not fully adult like in their reading and need to make regressions for comprehension, but the children with more detailed lexical representations are better able to judge when they are having comprehension difficulty and so are more likely to regress, which in turn suggests that they are more capable of modulating their eye-movement behavior to support successful comprehension than are their peers.

In the analyses comparing eye-movement behavior across the different tasks, the difference in fixation durations between reading and the other tasks was greatest for children with higher quality lexical representations. In other words, as comprehension abilities improve, the difference in eye-movement behavior between reading and other visual tasks increases, and reading becomes more specialized and adult like. This finding is important in that it shows that individual differences in reading skill reflect language-specific development and not (or at least,
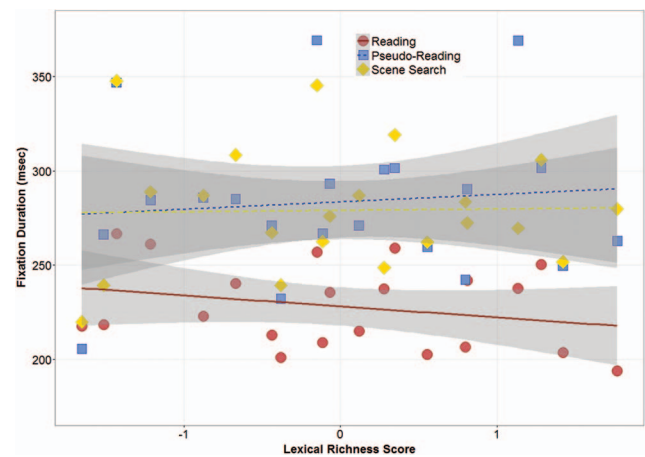

Figure 5. The interaction of task with lexical richness score in the analysis of fixation durations. Readers with higher quality lexical representations showed a greater divergence in fixation duration between reading and the other tasks. Error bands represent 95% confidence intervals. See the online article for the color version of this figure.

not just) improvements in eye-movement control as a result of task-independent oculomotor development. This is highly consistent with modeling work from the EZ reader model of eye-movement control, which implicates lexical processing as the primary determinant of reading differences between children and adults (Reichle et al., 2013).

Unlike lexical quality, decoding skill (as measured by speed and accuracy of oral reading of words and nonwords using the TOWRE) had almost no influence on eye movements in any task. This is perhaps not surprising considering that participants were screened to ensure that all were adequate decoders, although there was enough variability in the measure that differences could have arisen if they existed. Even so, this finding suggests that it is the quality of lexical-semantic representations, rather than more superficial orthographic or phonological ones, that underlie most individual differences among adolescent readers, particularly with young adolescents like the ones tested here.

Conclusion

This study investigated how individual differences in the quality of lexical representations influence eye-movement behavior in reading. Adolescents with higher quality lexical representations were generally more efficient in their eye movements, with shorter gaze durations and fewer refixations compared to their peers with lower quality lexical representations. Further, the influence of lexical quality was specific to reading, indicating that these individual differences reflected the influence of language development on reading skill and not just general oculomotor development. These results suggest that, as children's ability to read improves with age, children become increasingly more skilled at using the eye-movement system for the specific task of reading, and gradually treat text differently from other kinds of materials that are interpreted using eye movements. Thus, lexical quality is a significant predictor of reading ability, and differences in lexical quality are associated with increased specialization of the eye-movement system for the task of reading.

References

Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science, 23,* 201–207. http://dx.doi.org/10.1177/0963721414530145

Blythe, H. I., Häikiö, T., Bertam, R., Liversedge, S. P., & Hyönä, J. (2011). Reading disappearing text: Why do children refixate words? *Vision Research, 51,* 84–92. http://dx.doi.org/10.1016/j.visres.2010.10.003

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

Blythe, H. I., Liversedge, 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,* 1583–1591. http://dx.doi.org/10.1016/j.visres.2009.03.015

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 699–706. http://dx.doi.org/10.1037/0278-7393.32.4.699

Davies, M. (2009). The 385 + million word *Corpus of Contemporary American English* (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics, 14,* 159–190. http://dx.doi.org/10.1075/ijcl.14.2.02dav

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first- and second-language word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology, 66,* 843–863. http://dx.doi.org/10.1080/17470218.2012.720994

Fukushima, J., Hatta, T., & Fukushima, K. (2000). Development of voluntary control of saccadic eye movements: I. Age-related changes in normal children. *Brain & Development, 22,* 173–180. http://dx.doi.org/10.1016/S0387-7604(00)00101-7

Hamilton, S. T., Freed, E. M., & Long, D. L. (2013). Modeling reader and text interactions during narrative comprehension: A test of the lexical quality hypothesis. *Discourse Processes, 50,* 139–163. http://dx.doi.org/10.1080/0163853X.2012.742001

Henderson, J. M., & Luke, S. G. (2012). Oculomotor inhibition of return in normal and mindless reading. *Psychonomic Bulletin & Review, 19,* 1101–1107. http://dx.doi.org/10.3758/s13423-012-0274-2

Henderson, J. M., & Luke, S. G. (2014). Stable individual differences in saccadic eye movements during reading, pseudoreading, scene viewing, and scene search. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 1390–1400. http://dx.doi.org/10.1037/a0036330

Joseph, H. S., 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. http://dx.doi.org/10.1016/j.visres.2009.05.015

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,* 207–222.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354. http://dx.doi.org/10.1037/0033-295X.87.4.329

Kirkby, J. A., Blythe, H. I., Drieghe, D., & Liversedge, S. P. (2011). Reading text increases binocular disparity in dyslexic children. *PLoS ONE, 6*(11), e27105. http://dx.doi.org/10.1371/journal.pone.0027105

Kirkby, J. A., Webster, L. A., Blythe, H. I., & Liversedge, S. P. (2008). Binocular coordination during reading and non-reading tasks. *Psychological Bulletin, 134,* 742–763. http://dx.doi.org/10.1037/a0012979

Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General, 135,* 12–35. http://dx.doi.org/10.1037/0096-3445.135.1.12

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. http://dx.doi.org/10.1016/j.jml.2011.03.002

Leech, R., Aydelott, J., Symons, G., Carnevale, J., & Dick, F. (2007). The development of sentence interpretation: Effects of perceptual, attentional and semantic interference. *Developmental Science, 10,* 794–813. http://dx.doi.org/10.1111/j.1467-7687.2007.00628.x

Luke, S. G., & Henderson, J. M. (2013). Oculomotor and cognitive control of eye movements in reading: Evidence from mindless reading. *Attention, Perception, & Psychophysics, 75,* 1230–1242. http://dx.doi.org/10.3758/s13414-013-0482-5

Luke, S. G., Nuthmann, A., & Henderson, J. M. (2013). Eye movement control in scene viewing and reading: Evidence from the stimulus onset delay paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 10–15. http://dx.doi.org/10.1037/a0030392

Luna, B., Velanova, K., & Geier, C. F. (2008). Development of eye-movement control. *Brain and Cognition, 68,* 293–308. http://dx.doi.org/10.1016/j.bandc.2008.08.019

Montgomery, J. W., Magimairaj, B. M., & O'Malley, M. H. (2008). Role of working memory in typically developing children's complex sentence

comprehension. *Journal of Psycholinguistic Research, 37,* 331–354. http://dx.doi.org/10.1007/s10936-008-9077-z

Nuthmann, A., Engbert, R., & Kliegl, R. (2007). The IOVP effect in mindless reading: Experiment and modeling. *Vision Research, 47,* 990–1002. http://dx.doi.org/10.1016/j.visres.2006.11.005

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11,* 357–383. http://dx.doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. *Precursors of functional literacy, 11,* 67–86.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41,* 211–236. http://dx.doi.org/10.1016/0022-0965(86)90037-8

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422. http://dx.doi.org/10.1037/0033-2909.124.3.372

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,* 1457–1506. http://dx.doi.org/10.1080/17470210902816461

Rayner, K., & Fischer, M. H. (1996). Mindless reading revisited: Eye movements during reading and scanning are different. *Perception & Psychophysics, 58,* 734–747. http://dx.doi.org/10.3758/BF03213106

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology: Section A, Human Experimental Psychology, 53,* 1061–1080. http://dx.doi.org/10.1080/713755934

Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research, 47,* 2714–2726. http://dx.doi.org/10.1016/j.visres.2007.05.007

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill.

Developmental Review, 33, 110–149. http://dx.doi.org/10.1016/j.dr.2013.03.001

Salman, M. S., Sharpe, J. A., Eizenman, M., Lillakas, L., Westall, C., To, T., . . . Steinbach, M. J. (2006). Saccades in children. *Vision Research, 46* (8–9),1432–1439. http://dx.doi.org/10.1016/j.visres.2005.06.011

Smith, C. L., & Tager-Flusberg, H. (1982). Metalinguistic awareness and language development. *Journal of Experimental Child Psychology, 34,* 449–468. http://dx.doi.org/10.1016/0022-0965(82)90071-6

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General, 136,* 64–81. http://dx.doi.org/10.1037/0096-3445.136.1.64

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency (TOWRE).* Austin, TX: Pro-Ed.

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science, 11,* 84–97. http://dx.doi.org/10.1111/j.1467-7687.2007.00656.x

Verhaeghen, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology and Aging, 18,* 332–339. http://dx.doi.org/10.1037/0882-7974.18.2.332

Vitu, F., O'Regan, J. K., Inhoff, A. W., & Topolski, R. (1995). Mindless reading: Eye-movement characteristics are similar in scanning letter strings and reading texts. *Perception & Psychophysics, 57,* 352–364. http://dx.doi.org/10.3758/BF03213060

Wassenberg, R., Hurks, P. P., Hendriksen, J. G., Feron, F. J., Meijs, C. J., Vles, J. S., & Jolles, J. (2008). Age-related improvement in complex language comprehension: Results of a cross-sectional study with 361 children aged 5 to 15. *Journal of Clinical and Experimental Neuropsychology, 30,* 435–448. http://dx.doi.org/10.1080/13803390701523091

Woodcock, R. (2011). *Woodcock Reading Mastery Tests—Third Edition (WRMT–R),* Circle Pines, MN: American Guidance Service.