Journal of Experimental Psychology:
Learning, Memory, and Cognition
2018, Vol. 44, No. 7, 1051–1063

0278-7393/18/$12.00 http://dx.doi.org/10.1037/xlm0000506

The Development of Wrap-Up Processes in Text Reading: A Study of Children's Eye Movements

Simon P. Tiffin-Richards and Sascha Schroeder
Max Planck Institute for Human Development, Berlin, Germany

Reading comprehension is the product of constructing a coherent mental model of a text. Although some of the processes that are necessary to construct such a mental model are executed incrementally, others are deferred to the end of the clause or sentence, where integration processing is wrapped up before the reader progresses further in the text. In this longitudinal study of 65 German-speaking children across Grades 2, 3, and 4, we investigated the development of wrap-up processes at clause and sentence boundaries by tracking the children's eye movements while they read age-appropriate texts. Our central finding was that children in Grade 2 showed strong wrap-up effects that then slowly decreased across school grades. Children in Grades 3 and 4 also increasingly used clause and sentence boundaries to initiate regressions and rereading. Finally, children in Grade 2 were shown to be significantly disrupted in their reading at line breaks, which are inherent in continuous text. This disruption decreased as the children progressed to Grades 3 and 4. Overall, our results show that children exhibit an adultlike pattern of wrap-up effects by the time they reach Grade 4. We discuss this developmental trajectory in relation to models of text processing and mechanisms of eye-movement control.

Keywords: eye movements, wrap-up effects, reading comprehension, reading development, children

Reading comprehension involves the construction of a mental model in which individual words are integrated and combined with relevant world knowledge (van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). Although some of these processes are incremental, clause and sentence endings have also been shown to be important for integration processes (Aaronson & Ferres, 1984). Specifically, readers pause at clause and sentence boundaries to "wrap up" new information and integrate it into the mental model (Just & Carpenter, 1980; Rayner, Kambe, & Duffy, 2000). The extent of wrap-up processing has been shown to be influenced by the information density of a sentence (Haberlandt & Graesser, 1985) and readers' processing capacity (Stine-Morrow et al., 2010). However, wrap-up processes are also affected by factors unrelated to integration, factors such as the presence of punctuation marks and the intonation pattern at clause boundaries (Hirotani, Frazier, & Rayner, 2006), and may also act as a strategic mechanism to avoid returning to a clause or sentence once a reader has moved on in the text (Warren, White, & Reichle, 2009).

Previous studies have shown that skilled adult readers pause at clause and sentence boundaries and that the factors that prompt these pauses may vary according to the characteristics of both the text and the reader. However, it is an open question whether beginning readers also engage in wrap-up processes, because they still struggle with word recognition and their processing is more localized than in adults. To address this question, we conducted a longitudinal study in which we tracked the eye movements of a group of children across school Grades 2, 3, and 4 while reading age-appropriate texts and compared them with the eye movements of adults reading the same materials. We were particularly interested in children's wrap-up processing at clause and sentence boundaries and how they change during reading development, but we were also interested in disruptions caused by line breaks inherent in connected text.

Wrap-Up Effects in Reading

There is a large body of evidence from eye-movement studies that indicates that adult readers spend more time at clause- and sentence-final words than nonfinal words (Just & Carpenter, 1980; Rayner et al., 2000; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989; Stine-Morrow et al., 2010). Rayner et al. (2000), for example, found longer gaze durations on clause-final nouns than on the same nouns in nonfinal clause positions. In addition, the likelihood of skipping a noun in clause-final position was smaller, and the likelihood of initiating a regression was greater, than for nouns in a nonfinal position (see also Rayner et al., 1989). Syntactic boundaries therefore appear to trigger the allocation of additional attentional resources, resulting in longer reading times, and may also initiate rereading of earlier sections of the text.

Wrap-up effects have traditionally been associated with integration processes that are important for the construction of the situation model of the text (Just & Carpenter, 1980). The process of constructing a coherent mental model is necessarily incremental because readers revise their situation model continuously as they progress through a text. Events unfold, new information is introduced, previous ambiguities are resolved, and new situational

This article was published Online First March 5, 2018.

Simon P. Tiffin-Richards and Sascha Schroeder, Reading Education and Development, Max Planck Institute for Human Development, Berlin, Germany.

Correspondence concerning this article should be addressed to Simon P. Tiffin-Richards, Reading Education and Development, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail: tiffin-richards@mpib-berlin.mpg.de

discontinuities are encountered. Hence, the situation model is not static but is updated as new information is added and old information is elaborated (Zwaan, Langston, & Graesser, 1995; Zwaan & Radvansky, 1998). Clearly, the incremental updating of the situation model requires significant cognitive resources to keep information active and to integrate new information into the mental representation. Because resources for discourse processing are limited, it has been suggested that readers engage in these processes cyclically at syntactic boundaries. According to this buffer–integrate–purge view, text information is extracted and buffered until the end of a clause or sentence, where it is integrated into a more durable text representation, upon which the buffer is purged to prepare for the next input (Haberlandt & Graesser, 1989; Haberlandt, Graesser, Schneider, & Kiely, 1986). Indeed, the end of a sentence often clarifies ambiguities and signals the conclusion of a train of thought, making this a good location for integration and comprehension processes (Just & Carpenter, 1980).

The special status of syntactic boundaries for initiating wrap-up does not, however, preclude incremental and ongoing information integration as the reader progresses through a text. Rayner, Warren, Juhasz, and Liversedge (2004), for instance, showed that readers are immediately sensitive to semantic violations in sentence reading. Reading times were immediately affected when an anomalous word was encountered, which suggests that the situation model is incrementally constructed as each word is read. Coregistration of eye movements and event-related brain potentials has also shown immediate responses to sematic and syntactic violations in sentence reading (Metzner, von der Malsburg, Vasishth, & Rösler, 2016). Integrational processing thus appears to be performed both online while reading the sentence and periodically at syntactic boundaries. It is unclear, however, what the division of labor between these two kinds of processes is and whether they differ between different groups of readers.

It is well known that wrap-up at syntactic boundaries is influenced by attributes of both the text and the reader. Readers can, for instance, be encouraged to perform wrap-up by manipulating the prominence (i.e., explicit punctuation) of syntactic boundaries. Hirotani et al. (2006) found that clause-final wrap-up was stable in fixation durations only when clauses were punctuated with a comma. Readers showed longer gaze durations at clause-final and sentence-final words than nonfinal words. When a clause did not end with a comma, there was no increase in gaze duration. One interpretation of this finding was that wrap-up has more to do with a reaction to punctuation than to the syntactic boundary (Hirotani et al., 2006). The wrap-up effect was also evident for prosodic boundaries at which a speaker would naturally pause but the need for interpretive processing was negligible (e.g., "*John,* go to the library for me"). These findings are evidence that readers use implicit intonation patterns to guide their eye movements, resulting in pauses at syntactic boundaries.

Wrap-up processes may also to some extent be strategic and self-regulated. In an eye-tracking study, Stine-Morrow et al. (2010) manipulated the prominence of a syntactic boundary (unmarked, comma-marked, full-stop-marked) in short text passages and measured effects at the manipulated boundary and the following sentence boundary. There were no differences between conditions in early reading measures at the clause boundary, although there were effects in the rereading time of a clause before moving past the clause boundary. More rereading was initiated at the boundary if it was comma- or full-stop-marked than when it was unmarked. It is important to note that Stine-Morrow et al. (2010) also found that in the unmarked condition, wrap-up increased at the subsequent sentence boundary. This was taken as evidence for a pay-now-or-pay-later principle, in which intermediate clause-final wrap-up facilitates processing of subsequent text. This account is consistent with the buffer–integrate–purge cycle model, because unresolved ambiguities or unintegrated information remains in the buffer and must be processed and purged at a later point in the text. It is also compatible with the view that syntactic integration is incremental across a text, in that periodic wrap-up provides better context information on which expectations of upcoming discourse can be based. It is interesting that Hirotani et al. (2006) also found that readers generally avoid crossing clause boundaries on regressions after first-pass reading, indicating that wrap-up processes may function as a strategy to avoid having to reread a clause after proceeding to the next clause or sentence.

What is particularly intriguing about the Stine-Morrow et al. (2010) study is that there were developmental differences in wrap-up processes between participants in different age groups. Older adult readers tended to be more sensitive to the salience of clause boundaries in early measures, suggesting that they were engaging in more frequent wrap-up than were the younger readers. This was interpreted as a compensatory strategy, in which older adults rely on more regular consolidation of their discourse model to counteract their less efficient lexical processing and working memory capacities. The compensatory interpretation is consistent with findings that older adults tend to rely on story structure and context when reading for comprehension to compensate for other processing deficits (Stine-Morrow, Soederberg Miller, Gagne, & Hertzog, 2008; Stine-Morrow, Soederberg Miller, & Leno, 2001).

Another important structural characteristic of continuous text is the necessity of line breaks, which typically do not coincide with clause or sentence boundaries and interrupt the natural flow of reading. However, only a few studies have investigated the effects of line breaks on adult readers. Rayner (1977), for instance, found that the last fixation on a line was generally shorter than nonfinal fixations. Kuperman, Dambacher, Nuthmann, and Kliegl (2010) similarly found a decrease in fixation durations toward the end of lines in multiline texts. One explanation provided for the decrease in fixation duration at the end of a line is that the processing of the last word cannot be influenced by parafoveal processing of an upcoming word (Rayner, 1977). Alternatively, Kuperman et al. (2010) and Mitchell, Shen, Green, and Hodgson (2008) suggested that end-of-line effects may also be due to oculomotor programming. It is important to note that in single-sentence studies there is no disambiguation between sentence-final and line-final processes, because these necessarily coincide. This is relevant because processes may work in opposite directions, reducing the chance of finding clear effects of either line-final speed-up or sentence-final wrap-up.

Children's Eye Movements and Integrational Processing

Children are clearly slower readers than are adults. They typically refixate words multiple times and make more regressions back to words they have already read, leading to extended first-pass and rereading times (Blythe, 2014; Schroeder, Hyönä, & Liversedge, 2015). Differences between the eye movements of

beginning readers and skilled adult readers dissipate with chronological age and reading ability (Blythe, Häikiö, Bertam, Liversedge, & Hyönä, 2011; Häikiö, Bertram, Hyönä, & Niemi, 2009; Huestegge, Radach, Corbic, & Huestegge, 2009), suggesting that the main difference between beginning and skilled reading lies in the efficiency of lexical and postlexical processing rather than any qualitative difference in reading processes between children and adults (Reichle et al., 2013).

However, because most studies conducted with children have used highly constrained single-sentence stimuli, how beginning readers' eye movements might be affected by properties of complex syntactic structures in longer texts remains largely unknown. The studies that have investigated reading time measures in young readers' processing of text generally focus on children in upper primary and secondary schools (e.g., de Leeuw, Segers, & Verhoeven, 2016; van der Schoot, Vasbinder, Horsley, & van Lieshout, 2008; van Silfhout, Evers-Vermeul, Mak, & Sanders, 2014). Nevertheless, a few eye-tracking studies with younger children have shed light on how children integrate information within sentences. Joseph et al. (2008), for instance, showed that semantic anomalies (e.g., "John used the pump to inflate the carrots for dinner") caused immediate disruption and longer first-pass reading times on the target word (e.g., *carrots*) in children, as previously found in adult readers (Rayner et al., 2004). This suggests that children incrementally update their situation model and are thus capable of detecting semantic violations as soon as they are encountered. However, if the violation was not semantically illegal but only implausible (e.g., "John used the axe to chop the carrots for dinner"), the effect was delayed in children and the disruption was observed in only the posttarget (and wrap-up) region, whereas adults still showed immediate effects on the target word *carrots.* This suggests that whereas children do react to semantic violations immediately, reactions to pragmatic inconsistencies may be delayed. This was also evident in a study conducted by Joseph and Liversedge (2013) in which children and adults read ambiguous sentences such as "The boy poked the elephant with the long stick/trunk from outside the cage." When the disambiguating noun *stick/trunk* was attached to the noun phrase (i.e., *trunk–elephant*)—which is the less preferred option in English—reading was disrupted for both adults and children compared to when it was attached to the verb phrase (i.e., *stick–poke*). The disruption was less immediate for children compared to adults, suggesting a delayed response to the initial incorrect attachment. Similarly, Joseph, Bremner, Liversedge, and Nation (2015) demonstrated that children's resolution of noun phrase anaphors (e.g., *vehicle*) was affected by both the distance between referent and anaphor and the typicality of the referent–anaphor association (e.g., when *vehicle* referred to *truck* [typical] compared to *crane* [atypical]). However, children did not resolve the anaphors when the referent was distant and atypical. Again, this suggests that children are sensitive to semantic features but that their processing is resource-intensive and delayed.

Taken together, the studies investigating children's integration processes suggest two things. First, children are able to process information incrementally as far as they are sensitive to anomalies, syntactic misalignments, and the typicality of referent–anaphor associations, generally as soon as these are encountered. Second, discrepancies that were more difficult to detect (such as pragmatic inconsistencies or less preferred syntactic attachment) result in either delayed processing or failure to detect the discrepancy at all.

The Present Study

To investigate how beginning readers respond to syntactic boundaries and line breaks, we conducted a longitudinal study in which the eye movements of German-speaking children were tracked from Grade 2 to Grade 4. In addition, we also collected eye-movement data from a sample of young adults reading the same materials. We were therefore able to investigate how the same materials are processed by readers of increasing reading ability. However, it is important to note that the data collected in the adult group are not directly comparable to that in earlier adult studies that used age-appropriate reading materials.

A further important point is that the present study was conducted in German, which differs in a few distinctive ways from English in its syntax. Because German makes intensive use of case marking, word order is generally quite flexible. In main clauses, the verb is generally located at the second position, whereas it takes the sentence-final position in subordinate clauses, unlike in English. Because the verb is critical for comprehension, this may increase reliance on sentence-final integration processes. This makes German a particularly interesting language in which to study wrap-up effects. In addition, in contrast to the case in English, punctuation rules are fairly systematic in German. Specifically, subordinate clauses are always separated from the main clause by a comma, but coordinated main clauses are generally not. This natural manipulation thus allowed us to investigate the effects of punctuation on clause wrap-up.

We had three main research questions. First, we investigated how wrap-up effects change during reading development. We predicted wrap-up effects would decrease longitudinally from Grade 2 to Grade 4 as children's reading abilities increase in efficiency and they become better at processing sentences incrementally. We also tested a number of factors presumed to moderate wrap-up processes, including clause punctuation, clause type, and clause and sentence length. To do this we used the natural occurrence of comma-marked and unmarked clause boundaries to check whether these affect wrap-up processes differently. Second, because children make more regressions while reading than do adults, we were interested in whether beginning readers use syntactic boundaries to initiate regressions in order to reread sections of a sentence and resolve comprehension difficulties before moving on in the text. We anticipated that children would increasingly rely on syntactic boundaries to initiate regressions with increasing reading ability and gradually make fewer regressions from within a clause or sentence. A final and more exploratory aim of our study was to investigate how line breaks affect beginning readers' eye movements. Because parafoveal processing develops with age and reading experience (Häikiö et al., 2009), we did not necessarily expect beginning readers to show the speed-up effects observed for adults. On the contrary, because line breaks are visually very salient, children may instead use them to trigger wrap-up processing, resulting in longer reading times.

Method

Participants

A total of 92 children from six Grade 2 classes at two primary schools in Berlin took part in the eye-tracking experiment as part

of a longitudinal study (the Berlin Developmental Eye Tracking Study). Of these, 65 children (71%) completed at least two eye-tracking sessions without technical problems and with comprehension scores on the reading materials above chance level. Of the effective longitudinal sample of 65 children (age: $M = 7$ years 9.5 months, $SD = 6$ months; 37 girls), 55 spoke only German at home and 10 spoke German and an additional language at home. A further 24 adults (age: $M = 25.75$, $SD = 3.23$, years; 14 female) were recruited using the participant database of the Max Planck Institute for Human Development, in Berlin. One adult was excluded from the analyses due to comprehension scores in the eye-tracking session below chance level. All adult participants were German speakers, had normal or corrected vision, and had no record of a reading disability. Adult participants provided written consent and were compensated with €10. Children participated with the written consent of their parents and school authorities and received a small gift after each test session. The children completed a standardized reading fluency test each year (the revised Salzburg Reading and Spelling Test [SRST-II]; Moll & Landerl, 2010), and adults completed the same test. Participants' T-scores for word fluency in Grade 2 ($M = 50$, $SD = 32$), Grade 3 ($M = 52$, $SD = 30$), and Grade 4 ($M = 60$, $SD = 25$) did not differ from the population mean ($M = 50$) of their respective age groups (all $ts < 2$). The adults' T-scores for word fluency ($M = 51$, $SD = 27$) also did not differ significantly from their population mean. Our participants were therefore within the normal range of reading ability. Ethical approval for this study was obtained from the Max Planck Institute for Human Development research ethics committee and the responsible school authorities.

Materials

Stories. The 18 text excerpts used in this study were drawn from the German childLex corpus of children's literature (Schroeder, Würzner, Heister, Geyken, & Kliegl, 2015) and represented a range of age-appropriate narrative texts with themes such as sea voyages, animals at the zoo, and staying at home alone. Text length ranged between five and 16 sentences ($M = 9.94$) and between 86 and 111 words ($M = 99$). The average word length was 4.89 letters ($SD = .22$), and the mean log-transformed lemma frequency per million was 6.68 ($SD = 3.11$). Sentences were an average of 12 words ($SD = 5$) long, and each line of text contained a maximum of 10 words. Each line of text was a maximum of 95 characters long, and line breaks occurred as necessary so as to place words on the screen without hyphenation.

Readability norms. General ratings of the stories' readability and age-appropriateness were collected from a sample of 67 Grade 2 children of similar age ($M = 7$ years 6 months, $SD = 5$ months) who did not participate in the eye-tracking study. Each child read three texts and rated their difficulty on a 4-point Likert scale ranging from 1 (*very difficult*) to 4 (*very easy*) and how much they enjoyed the story on a scale from 1 (*I did not enjoy it at all*) to 4 (*I enjoyed it very much*). The stories were rated on average as easy to understand ($M = 3.15$, $SD = .39$) and enjoyable to read ($M = 3.32$, $SD = .32$).

Boundary type. A boundary-type factor was used to code each word on its clause, sentence, and line boundary status. Words that did not end a clause, sentence, or line break were coded as nonboundary words. Words at the end of a clause, excluding sentence endings, were considered clause-final, whereas words at the end of a sentence were considered sentence-final. Finally, words at the end of a line of text, excluding cases in which this coincided with clause or sentence boundaries, were coded as line-final words. Words in different boundary positions differed in their average length and frequency. As displayed in Table 1, clause- and sentence-final words were on average longer, less frequent, and more likely to be a content word than were nonboundary and line-final words. Accordingly, the effects of boundary type were controlled for word length and word frequency in all analyses.

It is further important to note that the clause boundaries were punctuated according to German grammar rules. In German, clause boundaries between main and subordinate clauses are always comma-separated. However, main clauses connected with coordinating connectors such as *und* (English *and*) and *oder* (English *or*) typically omit the comma. Consequently, two thirds of the clause boundaries were comma-separated and one third were not.

Apparatus

An EyeLink 1000 eye tracker (SR Research, Kanata, Ontario, Canada) was used to record eye movements during reading at a rate of 1000 Hz. Text stimuli were presented on a 21-in. ASUS LCD monitor, with a refresh rate of 120 Hz. Participants sat at a viewing distance of 65 cm with an assisting head and chin rest to reduce head movements. Texts were presented in white in monospaced courier new font, size 14, on a black background using the UMass Eye Track 7.10m software (Stracuzzi & Kinsey, 2006).

Procedure

Testing took place within the children's schools in rooms suitable for eye tracking. In Grade 2 the testing was conducted during the second half of the school term. Testing in Grades 3 and 4 was conducted in roughly 12-month intervals. Adults were tested in a single session in laboratory rooms at the Max Planck Institute for Human Development, in Berlin. A nine-dot calibration of the eye tracker was conducted and validated with each participant until a calibration accuracy of at least .5° was achieved. The eye tracker was recalibrated after practice trials as well as after each break and as necessary when *x*- or *y*-axis drift was detected. Because the study was designed to investigate silent reading processes, children were instructed to read each story quietly to themselves and were

Table 1
Number of Words in Each Boundary Position and Their Average Word Length, Frequency, and Likelihood of Being a Content Word

		% Content words[a]		Word length[b]		Word frequency[c]	
Boundary type	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Nonfinal	1,358	.47	.50	4.60	2.11	7.01	3.11
Clause-final	79	.75	.44	5.67	2.19	5.61	2.92
Sentence-final	179	.83	.37	6.39	2.59	5.00	2.54
Line-final	166	.57	.50	5.20	2.48	6.32	3.08

[a] Included nouns, verbs, and adjectives. [b] Measured in number of characters. [c] Measured in log-transformed lemma frequency.

reminded whenever vocalizations were audible to the test instructor. Reading was binocular, and the right eye was tracked.

Texts were presented on two sequential screens in up to eight double-spaced lines. Each text was preceded by a fixation cross in the top left-hand corner of the screen that triggered the presentation of the trial automatically on fixation. Participants proceeded to the second screen of the text by pressing the *X* button on a gamepad. After reading the second screen of the text, participants ended the trial by again pressing the *X* button. The second screen of each text was then followed by three comprehension questions with four options each, labeled *A, B, C,* and *D.* Participants responded to each comprehension question by naming the label of the correct answer for the test instructor. After documenting the participant's response, the test instructor initiated the presentation of the next question. The 18 texts were divided into three lists of six trials. Participating children read six texts each, in random order. Lists were randomly assigned to participants in the order of their appearance in Grade 2. In Grades 3 and 4, children were reassigned to new lists so that each child read a different list each year. In six cases incorrect lists were assigned in Grade 4, and these data were omitted from the analyses. Adults read all three lists containing all 18 texts, in random order.

Analysis

Eye-movement data. Eye-movement data points were initially cleaned at the level of individual fixations. First, fixations of less than 80 ms were combined with an adjacent fixation if this was within one character. Shorter fixations of 40 ms or less were deleted if within three characters of the nearest fixation. Second, all remaining fixations under 60 ms or above 1,200 ms were discarded. This removed 1.6% of the children's fixations and 1.4% of the adults' fixations. Texts were not analyzed for children if more than half of the words in the text were skipped and the visualization of the eye movement pattern suggested skipping of large sections of text, indicating reading avoidance. This resulted in the exclusion of data from a total of six trials in Grade 2, six trials in Grade 3, and two trials in Grade 4. This behavior was not observed for adults.

Fixation measures. Four eye-movement measures were calculated from the fixation report (Rayner, 1998), including *first fixation* (all single and first of multiple fixations on a word), *gaze duration* (sum of all fixations on a word before the first saccade leaves the word), *first-pass regression probability* (likelihood of a regression out of a word on the first pass), and *regression-path duration* (sum of all fixations from the first fixation on a word up to but not including the first fixation to the right of the word, also referred to as go-past time). Regression-path duration is described as the time spent rereading text upon encountering a processing problem and before moving on to new text information. Unlike first fixation and gaze duration, this measure thus provides temporal information about the time course of processing disruptions and possible repair patterns consisting of regressions and rereading (see Liversedge, Paterson, & Pickering, 1998, for a more detailed discussion). In a final model criticism step, we deleted data points with residuals 2.5 *SD*s above the word and subject-specific means for each eye-movement measure (Baayen & Milin, 2010). In this final cleaning step, less than 2% of eye-movement data points were deleted for children and adults in each dependent measure. The entire cleaning procedure was conducted separately for adults and for children at each grade level.

Mixed-effects analysis. Linear-mixed-effects models were used to analyze the eye-movement data for each eye-movement measure in the R environment (R Development Core Team, 2015) with the lme4 package, Version 1.7 (Bates, Maechler, Bolker, & Walker, 2015). Participants and words were treated as crossed random effects, and all fixation-duration measures were log-transformed. Age group was included as a fixed effect with four levels (Grade 2, Grade 3, Grade 4, adults), which resulted in a repeated-measures design for children and allowed the direct cross-sectional comparison with skilled adult readers. Contrasts were used to compare each level of age group with the next highest level. Boundary type was included as a within-subject fixed effect and was effect-coded. Contrasts were used to compare words in the baseline nonfinal position with words in clause-final, sentence-final, and line-final positions. Words' length and log-transformed frequency were centered and included as continuous control variables. Cell-mean coding was used and planned contrasts were estimated for each effect using the multcomp R package function glht (Hothorn, Bretz, & Westfall, 2008). For fixation measures, contrasts were back-transformed from their logarithmic model estimates and are reported in milliseconds. For regression probability, contrasts were back-transformed from their logit model estimates and reported as probabilities.

The random effects structure was tested by including random intercepts for the variables subject, word, line, clause, sentence, and story and progressively excluding the random effects explaining the least amount of variability in the model. Each reduced model was tested against the previous model using a chi-square test of model fit with the anova function in R. The only random effects to significantly reduce model fit if removed were the random intercepts for subjects and words, which were hence included in all following analyses.

Results

The global eye-movement measures displayed in Table 2 show a typical developmental pattern for children from Grade 2 to Grade 4 (Huestegge et al., 2009) and cross-sectional comparison with skilled adult readers (Blythe, 2014; Schroeder et al., 2015). With increased reading experience, children made progressively longer saccades and fewer fixations, refixated words less often, and were increasingly likely to skip words. Overall regression probability, however, remained stable across grade levels. In contrast, skilled adult readers made fewer fixations overall, refixated words less often, and were more likely to skip words than were children. Adults also showed longer forward saccades and were less likely to initiate regressions on first-pass than were children of all grade levels. Comprehension scores on the texts were generally high and increased significantly from Grades 2 to 3 ($t > 2$) but did not differ between children in Grades 3 and 4 ($t < 2$). Children's comprehension scores in Grades 3 and 4 did not differ significantly from those of adults (all ts < 2). This indicates that although beginning readers in Grade 2 had poorer overall comprehension, by Grade 3 they were similar to skilled adult readers in being able to derive an understanding of the age-appropriate narrative texts, albeit at a much slower rate.

Table 2
Global Eye-Movement Measures and Comprehension Scores of Children and Adults

Measure	Grade 2		Grade 3		Grade 4		Adults	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Forward saccade (characters)	6.43	.05	7.19	.04	7.36	.04	9.97	.07
First-pass fixations	1.87	.01	1.50	.00	1.34	.00	1.14	.00
All fixations	2.68	.01	2.14	.01	1.83	.01	1.55	.01
Refixation probability	.47	.00	.35	.00	.28	.00	.14	.00
Skipping probability	.05	.00	.09	.00	.11	.00	.25	.00
Regression probability	.22	.00	.23	.00	.22	.00	.17	.00
First-fixation duration	274	.78	244	.64	224	.54	193	.37
Gaze duration	524	2.31	361	1.27	296	.90	217	.51
Regression-path duration	933	9.72	616	4.78	465	3.23	338	3.65
Comprehension score (%)	79	1.74	89	1.03	91	.82	93	.67

Note. Saccades from the end of each line of text were excluded from the mean saccade lengths for children and adults.

Wrap-Up Effects

The main effects and interactions of age group, word length and frequency, and boundary type on first fixations, gaze duration, regression-path duration, and regression probability are displayed in Table 3. The effects of word length in characters and log-transformed lemma frequency were included to control for differences in word-level properties and are not reported in the following sections. The main effect of age group was significant in all analyses, in that children had progressively shorter first fixations, gaze durations, and regression-path durations, as well as lower regression probability across grade levels. Adults had consistently shorter fixation durations in all measures and lower regression probability than did children of all grade levels. The main effect of age group is therefore omitted from the following sections, which focus on the interactions of age group and boundary type. The observed means for fixation-duration measures and regression probabilities for words in nonfinal, clause-final, sentence-final, and line-final position are displayed in Table 4 for children of all grade levels and adults. The ensuing wrap-up effects in fixation measures and regression probability at syntactic boundaries and speed-up effects at line breaks are reported in the text as back-transformed model estimates, controlled for the effects of word length and frequency for each age group. The coefficients, standard errors, and *p* values of the contrasts for the wrap-up effects are displayed in Table 5. The distributions of these effects are further illustrated in Figure 1.

First fixations. After controlling for the main effects of word length and frequency, as well as their interactions with age group, there was a significant main effect of boundary type on first-fixation durations and a significant interaction of boundary type and age group. At clause boundaries, there was no significant clause wrap-up effect for children or for adults in first-fixation duration, except for a small 7-ms wrap-up effect for children in Grade 3.

At sentence boundaries, however, there was a significant 12-ms sentence wrap-up effect for children in Grade 2, which remained stable at 12 ms in Grade 3 and 14 ms in Grade 4. The significant 6-ms sentence wrap-up for adults was significantly smaller ($b = .03$, $SE = .01$), $t(1) = 3.26$, $p = .001$, than for children in Grade 4.

At line boundaries, there was no effect in Grade 2. However, this changed significantly in Grade 3 ($b = .04$, $SE = .01$), $t(1) = 4.81$, $p < .001$, where children showed a significant line-final speed-up effect with first fixations 7 ms shorter at line-final words than nonfinal words. The line-final speed-up effect increased sig-

Table 3
Effects of Age Group, Boundary Type, and the Word Properties Length and Frequency on Fixation Measures and Regression Probability

Variable	First fixations			Gaze duration			Regression probability			Regression path		
	χ^2	*df*	*p*	χ^2	*df*	*p*	χ^2	*df*	*p*	χ^2	*df*	*p*
Intercept	188,806	1	<.001	58,254	1	<.001	491	1	<.001	37,758	1	<.001
Effects												
Age group	2,079	3	< .001	6,792	3	<.001	19	3	<.001	11,299	3	<.001
Boundary type	116	3	<.001	85	3	<.001	107	3	<.001	262	3	<.001
Boundary Type × Age Group	319	9	<.001	227	9	<.001	312	9	<.001	344	9	<.001
Control variables												
Word length	1	1	.536	1,117	1	<.001	5	1	.020	923	1	<.001
Word frequency	13	1	<.001	49	1	<.001	13	1	<.001	81	1	<.001
Word Length × Age Group	226	3	<.001	819	3	<.001	2	3	.569	732	3	<.001
Word Frequency × Age Group	7	3	.089	108	3	<.001	18	3	<.001	185	3	<.001

Table 4
Observed Means (and Standard Errors) of Fixation Measures and Regression Probability for Boundary and Nonboundary Words for Children and Adults

Word type	First fixations	Gaze duration	Regression probability	Regression-path duration
Grade 2				
Nonfinal	274 (1)	498 (2)	.20 (.002)	881 (10)
Clause-final	270 (3)	590 (12)	.25 (.011)	1,106 (47)
Sentence-final	282 (2)	665 (9)	.30 (.008)	1,342 (38)
Line-final	277 (2)	561 (8)	.28 (.008)	1,024 (28)
Grade 3				
Nonfinal	244 (1)	350 (1)	.21 (.003)	583 (5)
Clause-final	247 (3)	397 (6)	.26 (.011)	674 (18)
Sentence-final	256 (2)	442 (5)	.35 (.008)	944 (24)
Line-final	236 (2)	347 (4)	.24 (.007)	584 (14)
Grade 4				
Nonfinal	226 (1)	289 (1)	.20 (.003)	434 (3)
Clause-final	233 (3)	323 (4)	.26 (.011)	532 (17)
Sentence-final	245 (2)	359 (3)	.35 (.008)	692 (15)
Line-final	207 (2)	269 (3)	.21 (.007)	421 (10)
Adults				
Nonfinal	195 (1)	215 (1)	.15 (.002)	299 (3)
Clause-final	198 (2)	224 (3)	.18 (.010)	322 (10)
Sentence-final	211 (1)	241 (2)	.35 (.008)	658 (22)
Line-final	178 (1)	197 (2)	.12 (.006)	281 (7)

nificantly ($b = .07$, $SE = .01$), $t(1) = 6.77$, $p < .001$, in Grade 4 to 19 ms. Adults also showed a significant 25-ms line-final speed-up effect, which was significantly larger than for children in Grade 4 ($b = .04$, $SE = .01$), $t(1) = 4.58$, $p < .001$.

These results suggest that there were only small clause boundary wrap-up effects in the early measure of first-fixation duration. Sentence wrap-up was evident across grades for children and was smaller for skilled adult readers. Beginning readers in Grade 2 showed no line-final effects, which developed into a progressively larger line-final speed-up effect in Grades 3 and 4 (see Figure 1, top row, right panel). Skilled adult readers showed a larger line-final speed-up effect in first-fixation durations than did children in Grade 4, suggesting a developmental trajectory toward increased reading speed at line-final words.

Table 5
Planned Contrast Estimates of Clause, Sentence, and Line Boundary Wrap-Up Effects in Fixation Measures and Regression Probability for Children and Adults

	Clause wrap-up				Sentence wrap-up				Line wrap-up			
Variable	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*	*b*	*SE*	*t*	*p*
First fixations												
Grade 2	$.012_a$	.014	.874	.382	$.049_a$	.010	4.917	<.001	$.019_a$	.010	1.903	.057
Grade 3	$.032_a$	.014	2.267	.023	$.053_a$	.010	5.331	<.001	$-.028_b$	.010	−2.830	.005
Grade 4	$.017_a$	.014	1.152	.249	$.063_a$	.010	6.215	<.001	$-.098_c$	.010	−9.547	<.001
Adults	$-.006_a$	.015	−.418	.676	$.029_b$	.010	2.934	.003	$-.144_d$	.010	−14.045	<.001
Gaze duration												
Grade 2	$.033_a$	.022	1.501	.133	$.091_a$	.015	5.881	<.001	$-.008_a$	.015	−.540	.590
Grade 3	$.034_a$	.022	1.532	.125	$.070_a$	.015	4.549	<.001	$-.088_b$	.015	−5.687	<.001
Grade 4	$.031_a$	.022	1.411	.158	$.083_a$	.016	5.271	<.001	$-.143_c$	.016	−9.080	<.001
Adults	$-.016_b$	.022	−.730	.465	$.020_b$	.016	1.316	.188	$-.162_c$	.016	−10.310	<.001
Regression probability												
Grade 2	$.233_a$	.145	1.613	.107	$.553_a$	.101	5.475	<.001	$.578_a$	.101	5.713	<.001
Grade 3	$.338_a$	.144	2.337	.019	$.829_b$	.100	8.290	<.001	$.283_b$	.102	2.775	.006
Grade 4	$.350_a$	.145	2.416	.016	$.866_b$	.101	8.532	<.001	$.168_b$	.104	1.613	.107
Adults	$.355_a$	.151	2.353	.019	1.489_c	.102	14.647	<.001	$.003_c$	.110	.030	.976
Regression-path duration												
Grade 2	$.061_a$	.030	2.036	.042	$.161_a$	.021	7.643	<.001	$-.034_a$	.021	−1.609	.108
Grade 3	$.075_a$	.030	2.505	.012	$.250_b$	.021	11.903	<.001	$-.130_b$	.021	−6.121	<.001
Grade 4	$.079_a$	.030	2.596	.009	$.255_b$	.021	11.893	<.001	$-.175_c$	.022	−8.123	<.001
Adults	$.021_b$	.030	.697	.486	$.378_c$	.021	17.825	<.001	$-.169_c$	.021	−7.897	<.001

Note. Clause wrap-up was calculated as a contrast between the model mean of nonfinal words and clause-final words. Sentence and line break contrasts were calculated analogously. Subscripts indicate differences in wrap-up effects between groups.

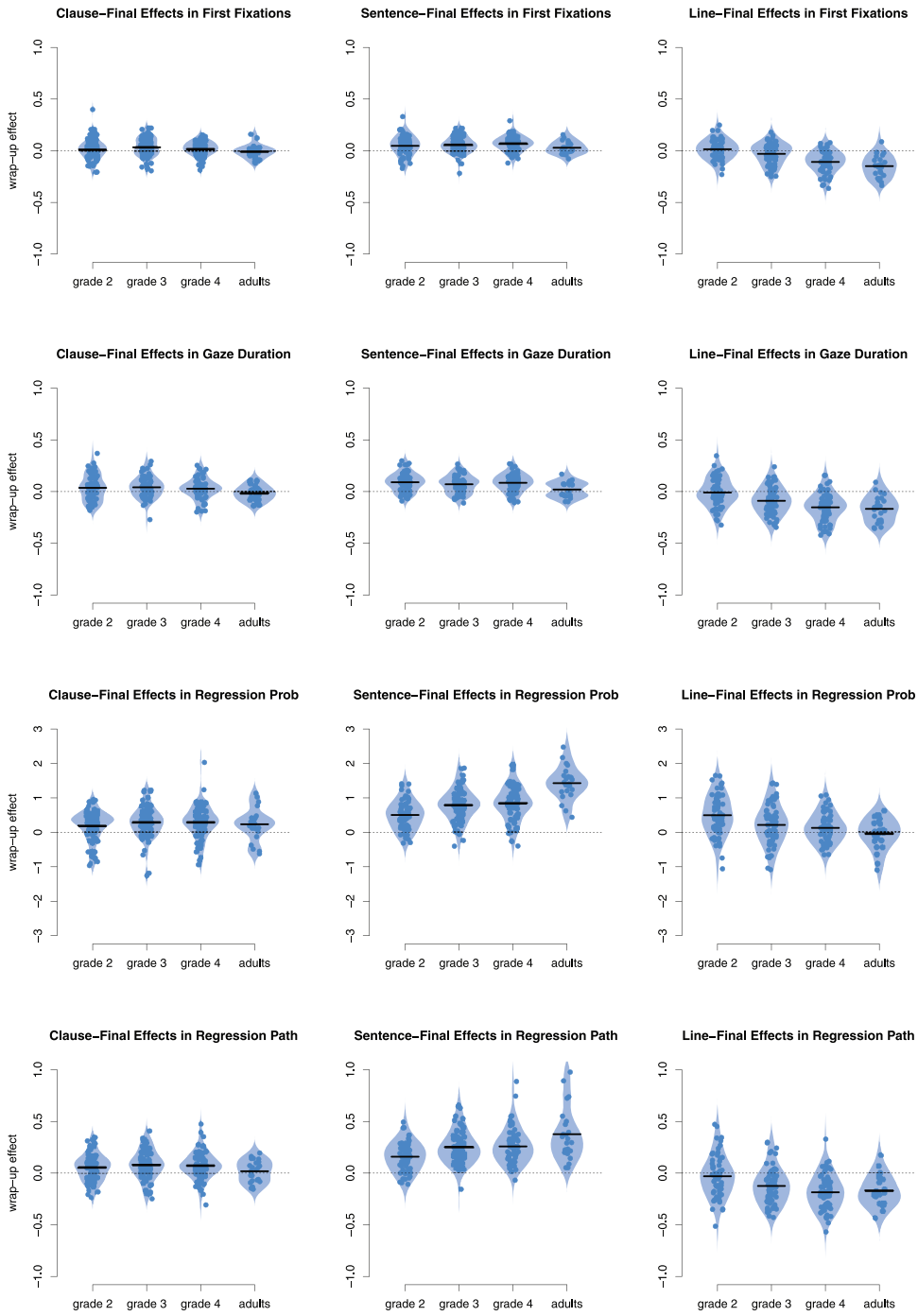

Figure 1. Wrap-up effect sizes at clause, sentence, and line boundaries in fixation measures and regression probability for children and adults. Effect sizes are reported as untransformed model estimates of mean fixation measures and regression probability (Prob) with their distributions and means. See the online article for the color version of this figure.

Gaze duration. After controlling for the main effects of word length and frequency, as well as their interactions with age group, we found a significant main effect of boundary type on gaze durations and a significant interaction of boundary type and age group. At clause boundaries, there were again no significant clause wrap-up effects for children or adults.

At sentence boundaries, however, there was a significant 38-ms sentence wrap-up effect for children in Grade 2. The sentence wrap-up effect did not change significantly with grade level, although the effect decreased numerically to 26 ms in Grade 3 and 22 ms in Grade 4. The sentence wrap-up effect was not significant for adults.

At line boundaries, beginning readers in Grade 2 did not show a line-final speed-up effect in gaze duration. This changed significantly in Grade 3 ($b = .07$, $SE = .01$), $t(1) = 6.65$, $p < .001$, at which point there was a significant 26-ms line-final speed-up effect. The line-final speed-up effect increased significantly ($b = .06$, $SE = .01$), $t(1) = 4.42$, $p < .001$, from Grade 3 to Grade 4 to 35 ms. Skilled adult readers showed a significant 30-ms line-final speed-up effect, which did not differ significantly from the speed-up effect in Grade 4 ($b = .02$, $SE = .01$), $t(1) = -1.49$, $p = .135$.

These results indicate that there were no wrap-up effects at clause boundaries in gaze duration (see Figure 1, second line, left panel). There was a significant sentence wrap-up effect for children across grade levels, which decreased numerically. Skilled adult readers did not show the wrap-up effect. Beginning readers in Grade 2 showed no effect of line breaks, but the line-final speed-up effect increased in strength from Grades 3 to 4 (see Figure 1, second row, right panel). At Grade 4, the speed-up effect was comparable in size with the effect found in skilled adult readers, suggesting a development toward a first-pass reading-speed facilitation at line-final words.

Regression probability. After controlling for the main effects of word length and frequency, as well as their interactions with age group, we found a significant main effect of boundary type on regression probability and a significant interaction of boundary type and age group. At clause boundaries, there was no increase in regression probability at clause-final words in Grade 2. However, there was a significant clause wrap-up effect in Grades 3 and 4, with an increase of regression probability at clause-final words of 5%. The significant 3% clause wrap-up effect for adults did not differ significantly from that of the children in Grade 4 ($b = .01$, $SE = .11$), $t(1) = .04$, $p = .965$.

At sentence boundaries, the significant 9% sentence wrap-up effect for beginning readers in Grade 2 increased significantly ($b = .27$, $SE = .07$), $t(1) = 4.21$, $p < .001$, in Grade 3 to 14% and remained stable at 15% in Grade 4. The significant 22% sentence wrap-up effect for adults was significantly greater than for the children in Grade 4 ($b = .62$, $SE = .07$), $t(1) = 9.19$, $p < .001$.

At line boundaries, children in Grade 2 showed a significant 9% line-final increase in regression probability, which decreased significantly ($b = .29$, $SE = .07$), $t(1) = 4.35$, $p < .001$, in Grade 3 to 4%. The line breaks had no effect on regression probability for children in Grade 4 or for adult readers.

The results suggest that wrap-up effects in regression probability at clause boundaries were stable from Grade 3 onward and comparable between children and skilled adult readers. The sentence wrap-up effect increased with grade level and was stronger for adults than for children, suggesting that the use of sentence endings to initiate regression increases with reading experience and proficiency (see Figure 1, third row, center panel). The increased regression probability at line breaks diminished across grade levels and was not apparent in skilled adult readers, suggesting that only inexperienced readers rely extensively on line breaks to cue rereading (see Figure 1, third row, right panel).

Regression-path duration. After controlling for the main effects of word length and frequency, as well as their interactions with age group, we found a significant main effect of boundary type on regression-path duration and a significant interaction of boundary type and age group. At clause boundaries, there was evidence of significant clause wrap-up effects in regression-path duration, unlike in the early fixation measures of first fixation and gaze duration. In Grade 2, there was a significant 38-ms clause wrap-up effect in regression-path duration. This effect did not change significantly with grade level, although the effects decreased numerically in Grade 3 to 34 ms and in Grade 4 to 28 ms. There was no significant clause wrap-up effect in regression-path duration for adults.

At sentence boundaries, beginning readers in Grade 2 also showed a significant 105-ms sentence wrap-up effect, which increased significantly ($b = .09$, $SE = .02$), $t(1) = 5.70$, $p < .001$, to 121 ms in Grade 3. In Grade 4 the sentence wrap-up effect did not change significantly in size, although it was numerically smaller at 99 ms. The significant 108-ms sentence wrap-up effect for adults was significantly greater than the wrap-up effect for children in Grade 4 ($b = .12$, $SE = .02$), $t(1) = 7.91$, $p < .001$.

At line boundaries, beginning readers in Grade 2 did not show an effect of line breaks in regression-path duration. However, there was a significant 52-ms line-final speed-up effect in Grade 3 and a significantly larger ($b = .05$, $SE = .02$), $t(1) = 2.77$, $p = .006$, line-final speed-up effect of 55 ms in Grade 4. The significant 37-ms line-final speed-up effect for skilled adult readers, although numerically smaller, did not differ significantly from the effect for children in Grade 4 ($b = .01$, $SE = .02$), $t(1) = .34$, $p = .74$.

In the present case, *regression-path duration* is defined as the reinspection time of words in a text before moving past a clause, sentence, or line boundary. Wrap-up effects in regression-path duration thus represent the use of specific positions in a text to initiate extensive rereading before progressing on to the next clause, sentence, or line. Our results suggest that clause wrap-up effects, although absent in the early measures of first fixation and gaze duration, were evident in the late measure of regression-path duration for children. Children showed stable clause wrap-up effects in regression-path duration across grade levels, whereas the effect was not evident in adults (see Figure 1, bottom row, left panel). Children also showed sentence wrap-up effects, which increased with grade level (see Figure 1, bottom row, center panel). The speed-up effect at line breaks increased across grade levels and was comparable between adults and children by Grade 4, suggesting a developmental trajectory toward an end-of-line reading-speed facilitation.

Moderation of Wrap-Up Effects

Clause punctuation. In separate analyses of the wrap-up effects at clause boundaries, we tested whether wrap-up differed between comma-marked and unmarked clause boundaries by adding an effect-coded punctuation factor into the model (comma-marked, unmarked). There were no significant interactions of punctuation and clause boundary, indicating no difference in the clause wrap-up effects for comma-marked and unmarked clauses. This finding is at odds with results reported in Hirotani et al. (2006) but may in part be due to the specific punctuation rules in German.

Clause type. In further analyses of the wrap-up effects at the end of main clauses and subordinate clauses, we ran a model with the fixed effects of age group, clause boundary (including those ending in a full stop), and clause type (main clause vs. subordinate clause). In first-fixation duration, wrap-up effects were greater at

clause-final words in a subordinate clause than a main clause for children in Grade 2 (Δ = 9 ms, b = .03, SE = .02), $t(1)$ = 1.82, p = .034, as well as in Grade 3 (Δ = 12 ms, b = .05, SE = .02), $t(1)$ = 2.72, p = .006, and Grade 4 (Δ = 8 ms, b = .04, SE = .02), $t(1)$ = 2.08, p = .02. There were no effects for adults. In gaze duration, wrap-up effects were greater at clause-final words in a subordinate clause than a main clause for children in Grade 2 (Δ = 38 ms, b = .09, SE = .03), $t(1)$ = 2.93, p = .002, but not for children in Grades 3 or 4 or for adults. There were no differences in wrap-up effects in regression probability or regression path between main clauses and subordinate clauses. These results suggest that in German sentences, beginning readers tend to engage in greater wrap-up at the end of subordinate clauses than main clauses, probably due to the common verb-final position in German subordinate clauses. Wrap-up at the end of clauses may therefore be driven more by their syntactic function and structure than by their punctuation.

Clause and sentence length. In a final set of complementary analyses, we added the fixed effects of clause and sentence length into separate mixed-effects models, as well as their interactions with clause and sentence wrap-up. Clause and sentence length were centered, and their effects on wrap-up processes were tested at ± 1 *SD*.

For clause boundaries, there were no effects of clause length in first fixations or gaze duration. There were no significant simple main effects of regression probability or regression-path duration. For sentence boundaries, there was a significant interaction of sentence length and sentence wrap-up in first fixations and three-way interactions with age group for gaze duration, regression probability, and regression-path duration. In first fixations, participants showed greater wrap-up effects at the end of long (+1 *SD;* 17 words) compared to short (−1 *SD;* seven words) sentences (Δ = 34 ms, b = .29, SE = .07), $t(1)$ = 4.26, $p < .001$. In gaze duration, wrap-up effects were greater in long sentences than short sentences for children in Grade 2 (Δ = 54 ms, b = .12, SE = .03), $t(1)$ = 4.23, $p < .001$, but not in Grades 3 and 4 or for adults. In regression probability, there was no effect of sentence length on wrap-up effects for children. However, for adults the sentence wrap-up effect increased with sentence length (Δ = 16%, b = .65, SE = .22), $t(1)$ = 2.98, p = .003. In regression-path duration, the sentence wrap-up effect increased with sentence length for children in Grade 2 (Δ = 107 ms, b = .13, SE = .04), $t(1)$ = 3.06, p = .002, and in Grade 3 (Δ = 88 ms, b = .14, SE = .04), $t(1)$ = 3.08, p = .002, as well as for adults (Δ = 121 ms, b = .32, SE = .04), $t(1)$ = 7.17, p = .002. The effect did not reach significance in Grade 4.

Taken together, sentence length influenced the expression of wrap-up effects in several eye-movement measures. Wrap-up increased for long sentences, as did the time spent rereading. This suggests that the amount of wrap-up processing readers engaged in was influenced by the amount of new information processed since the last syntactic boundary.

Discussion

The results of our study provide several novel findings that add to the understanding of the development of wrap-up processes while reading continuous text. It is important to note that although all age groups showed evidence of extensive wrap-up processing, the different eye-movement measures had different developmental trajectories. Our results can be grouped into three main findings. First, beginning readers showed strong wrap-up effects in their fixation measures, and the size of these effects decreased continuously across reading development. This suggests a greater reliance on clause- and sentence-final wrap-up processes in less proficient readers, presumably because they need more resources for lexical decoding and thus have to defer some integrational processing to clause and sentence endings.

Second, we also found a developmental trend toward the initiation of regressions from syntactic boundaries. Beginning readers in Grade 2 did not make more regressions from the end of a sentence than from nonfinal words. However, as children progressed from Grade 2 to Grades 3 and 4, the proportion of sentence- and clause-final regressions increased and became continuously more adultlike, with the majority of regressions being initiated from sentence boundaries rather than from words within a sentence. This suggests that children increasingly use syntactic boundaries to initiate rereading.

Our third finding is the exact opposite developmental trajectory at line breaks. Specifically, children in Grade 2 appeared to respond to line breaks in a fashion similar to that for syntactic boundaries, in that regression probability increased on the last word of each line. This pattern was strikingly different from that found in adult readers, who showed decreased fixation durations and no increases in regression probability at the end of a line. The effects at line breaks changed considerably across grade levels. In Grades 3 and 4, fixation measures and regression probability decreased at line boundaries, indicating the line-final speed-up effect also found in adult readers. By Grade 4, the children's reaction to line breaks was very similar to that of skilled adult readers.

Our results show clear developmental trends. Children's pattern of wrap-up processing became continuously more adultlike with increasing reading skill, and by Grade 4, the patterns of eye movements reached a high degree of similarity. The general pattern suggested that although wrap-up effects were initially expressed strongly in fixation duration, wrap-up was more evident in regression probability in adults reading the same text materials. In the following sections, we discuss the implications of these findings in more detail.

Clause and Sentence Wrap-Up

In this study, we investigated whether beginning readers use linguistic cues to prompt wrap-up processes while reading continuous texts. Our results clearly suggest that they do. In the middle of their second year of formal literacy education, children already used syntactic boundaries to initiate sentence wrap-up processes, despite having to refixate individual words multiple times and spending an extensive amount of time rereading words. Children showed larger sentence-final wrap-up effects than do adults in their gaze durations across grade levels, suggesting that children invest extensive attentional resources to integrate information at sentence boundaries.

Our study also provides evidence that wrap-up processes in beginning readers are, at least in part, related to information integration. For instance, we found wrap-up effects at clause boundaries, regardless of whether they were marked by a comma

or not. As has already been suggested, this may be due to the strict punctuation rules in German. We also found that subordinate clauses, which tend to use a verb-final position and thus delay important sentence information until the end of the clause, enhanced wrap-up effects. Also, sentence length increased wrap-up effects. Together, this suggests that the amount of information that needs to be held active increases the amount of effort required for beginning readers to construct their situation model, whereas more skilled readers appear to process the same texts more incrementally.

This leads to a more general point in the comparison of children and adult readers. Our results show that beginning readers rely more heavily on wrap-up processes than do skilled readers. There may, however, be a subtler relationship between text difficulty and reading ability. As previous studies have indicated (e.g., Joseph et al., 2015), children tend to disregard discrepancies if these are subtle or when resolution may be too costly when, for example, the point of disambiguation is too far back in the text. The relationship of text difficulty and reading ability could thus be conceptualized as an inverse U-shape function, where simple texts with relatively low demand compared to the reader's ability reduce the need for buffering and wrap-up processes. This is likely to be the case for the skilled participants in our study, because the texts did not present them with the necessary challenge to require information buffering and extensive wrap-up processing. Increasing text difficulty in relation to reading ability increases the need for wrap-up, until the point where difficulty exceeds ability to the extent that a reader forgoes wrap-up due to unsustainable processing costs. We see the investigation of this relationship of reading ability and text difficulty as a promising area for future studies, particularly from a developmental perspective.

Initiation of Rereading at Syntactic Boundaries

Our results further revealed that although children and adults showed a similar increase in regression probability at clause boundaries, beginning readers in Grade 2 showed a much smaller increase at sentence boundaries than did adults, despite adults' making fewer regressions overall. The increased likelihood of skilled adult readers' making a regression at a syntactic boundary has previously been explained as a component of the wrap-up process. Rayner et al. (2000) suggested that a trade-off takes place between the processing cost of regressing back to a previously fixated word to resolve a comprehension problem and the risk of having to make a longer regression later in the text if the problem remains unresolved. This is similar to the pay-now-or-pay-later mechanism suggested by Stine-Morrow et al. (2010), which predicts larger processing costs down the line if integration problems are not dealt with before a clause boundary is crossed. Eye-tracking studies have indeed suggested that adult readers are reluctant to regress back across clause and sentence boundaries once they have been passed (Hirotani et al., 2006; Stine-Morrow et al., 2010), which supports the hypothesis that skilled readers use syntactic boundaries to integrate information, monitor comprehension, and initiate regressions back in the same clause or sentence to resolve comprehension difficulties before progressing further in the text. Our results show that children increasingly use both clause and sentence boundaries to initiate regressions across grade levels. By Grade 4, children showed the same increase in regression probability at sentence boundaries as did skilled adult readers.

The stable clause wrap-up effect found in children's regression-path durations indicates that they consistently spent more time rereading sections of text after a regression from a clause boundary than from a nonfinal word. Hence, children appeared to use clause boundaries to initiate more extensive rereading of previous words before they moved on to the next clause. Adults did not spend this extra time rereading following regressions from clause boundaries, indicating that they processed the same clauses more incrementally. The sentence wrap-up effect in regression-path duration increased across grade levels. This suggests that when reanalysis of a sentence occurs, more skilled readers initiate rereading from sentence boundaries and then spend a considerable amount of time rereading words before moving on to the next sentence. This is again consistent with the pay-now-or-pay-later mechanism, according to which it is more efficient to address comprehension difficulties in a sentence before the sentence boundary is crossed (Stine-Morrow et al., 2010). Beginning readers, on the other hand, also initiated rereading from clause boundaries, suggesting a strategy of more regular resampling of previous words that required more frequent regressions. This may represent a deliberate rereading strategy to facilitate situation model updating and reduce the risk of longer regressive saccades back to a sentence after a reader has already progressed to the next sentence. Frequent regressions followed by rereading may also indicate an overload of the buffer, because beginning readers struggle to keep information active and available as they progress through a text. This interpretation is supported by our finding that sentence length increased rereading time.

Line-Final Speed-Up

A further interesting finding of our study is the development of the line-final speed-up effect. Skilled readers' first fixation and gaze durations clearly decreased at line breaks. Shorter final fixations on a line have previously been reported by Rayner (1977), as has a general decrease in fixation durations toward the end of a line of text (Kuperman et al., 2010). Rayner (1977) suggested that shorter fixations on line-final words may be due to the absence of a word to the right of the current fixation, eliminating the need to process parafoveal information of the upcoming word. Beginning readers in Grade 2 did not show the same decrease in fixation durations at line-final words. This finding is consistent with the assumption that the capacity of parafoveal processing develops with reading experience and proficiency (Häikiö et al., 2009; Marx, Hawelka, Schuster, & Hutzler, 2015; Pagán, Blythe, & Liversedge, 2016; Tiffin-Richards & Schroeder, 2015), which explains why beginning readers do not show the same reading-speed facilitation as do adults when no parafoveal information is available. The developmental trajectory of the line break effect supports this interpretation. Children in Grade 3 already showed a line-final speed-up effect, which increased in Grade 4 to an effect size comparable to that found in adults. Children at age 7 have been shown to have a smaller span of letters from which parafoveal information can be extracted. However, this span reaches an adultlike level at the age of 11, suggesting a steep increase in the use parafoveal information across the primary school years (Häikiö et al., 2009). Nevertheless, because we did not manipulate parafoveal

load directly in the current study, increased parafoveal processing is a tentative albeit plausible explanation for the increasing reading speed at line boundaries across school grades and between children and adults.

Line breaks also triggered more regressions in beginning readers in Grade 2 than in readers in Grades 3 and 4. This suggests that beginning readers use line breaks to initiate regressions to resolve comprehension problems, in contrast to skilled readers, who use the more appropriate clause and sentence boundary. This is consistent with the assumption that regressions are sometimes initiated to stall the input of new information, thus buying time to complete ongoing processing (Mitchell et al., 2008). This account is plausible if children are assumed to have lower lexical processing capacities and use a "time out" strategy (Mitchell et al., 2008) to spend more time integrating information. However, this is a viable strategy only if the rereading initiated at line boundaries does not result in reprocessing of information. Indeed, regression-path durations were shorter following regressions from line-final words than from nonfinal words, suggesting that regressions initiated from a line break did not generally lead to extensive rereading. The line-final speed-up effect increased across grade levels, suggesting a developmental trend toward children's not using line breaks to initiate regressions and, when they do, spending less time rereading previous words in the line of text before moving on.

Taken together, our results show that although beginning readers in Grade 2 use syntactic boundaries to initiate regressions and rereading, they do not profit from increased reading speed at line breaks and appear to use line breaks to trigger rereading. Visual line breaks hence significantly influenced beginning readers' reading processes in continuous text. This disruption decreased with reading experience and proficiency and disappeared by Grade 4.

Conclusion

Our study provides the first longitudinal eye-movement data on how children engage in wrap-up processes while reading continuous text for comprehension. Taken together, our data suggest that text reading involves a number of processes, which include both incremental processing and periodic wrap-up, which require the allocation of attentional resources. Decisions concerning where and when to invest resources in wrap-up are driven by an increasingly efficient trade-off between (a) investing time in integrating or initiating rereading to resolve comprehension problems and (b) reducing the risk of costly later regressions back to the current clause or sentence after a reader has already progressed further along in the text. Our eye-movement data suggest that children become increasingly efficient during primary school in allocating their cognitive resources to processing texts incrementally, in initiating rereading at strategic points at syntactic boundaries, and in avoiding the disruption caused by line breaks.

References

Aaronson, D., & Ferres, S. (1984). Reading strategies for children and adults: Some empirical evidence. *Journal of Verbal Learning and Verbal Behavior, 23,* 189–220. http://dx.doi.org/10.1016/S0022-5371(84)90137-3

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3,* 12–28. http://dx.doi.org/10.21500/20112084.807

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48.

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

de Leeuw, L., Segers, E., & Verhoeven, L. (2016). The effect of student-related and text-related characteristics on student's reading behaviour and text comprehension: An eye movement study. *Scientific Studies of Reading, 20,* 248–263. http://dx.doi.org/10.1080/10888438.2016.1146285

Haberlandt, K. F., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General, 114,* 357–374. http://dx.doi.org/10.1037/0096-3445.114.3.357

Haberlandt, K., & Graesser, A. C. (1989). Buffering new information during reading. *Discourse Processes, 12,* 479–494. http://dx.doi.org/10.1080/01638538909544741

Haberlandt, K. F., Graesser, A. C., Schneider, N. J., & Kiely, J. (1986). Effects of task and new arguments on word reading times. *Journal of Memory and Language, 25,* 314–322. http://dx.doi.org/10.1016/0749-596X(86)90004-5

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. http://dx.doi.org/10.1016/j.jecp.2008.04.002

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language, 54,* 425–443. http://dx.doi.org/10.1016/j.jml.2005.12.001

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal, 50,* 346–363.

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. http://dx.doi.org/10.1016/j.visres.2009.09.012

Joseph, H. S., Bremner, G., Liversedge, S. P., & Nation, K. (2015). Working memory, reading ability and the effects of distance and typicality on anaphor resolution in children. *Journal of Cognitive Psychology, 27,* 622–639.

Joseph, H. S., & Liversedge, S. P. (2013). Children's and adults' on-line processing of syntactically ambiguous sentences during reading. *PLoS ONE, 8,* e54141. http://dx.doi.org/10.1371/journal.pone.0054141

Joseph, H. S., 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. http://dx.doi.org/10.1080/17470210701400657

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

Kuperman, V., Dambacher, M., Nuthmann, A., & Kliegl, R. (2010). The effect of word position on eye-movements in sentence and paragraph reading. *Quarterly Journal of Experimental Psychology, 63,* 1838–1857. http://dx.doi.org/10.1080/17470211003602412

Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye movements in reading and scene perception* (pp. 55–75). Amsterdam, the Netherlands: Elsevier. http://dx.doi.org/10.1016/B978-008043361-5/50004-3

Marx, C., Hawelka, S., Schuster, S., & Hutzler, F. (2015). An incremental boundary study on parafoveal preprocessing in children reading aloud: Parafoveal masks overestimate the preview benefit. *Journal of Cognitive Psychology, 27,* 549–561. http://dx.doi.org/10.1080/20445911.2015.1008494

Metzner, P., von der Malsburg, T., Vasishth, S., & Rösler, F. (2017). The importance of reading naturally: Evidence from combined recordings of eye movements and electric brain potentials. *Cognitive science, 41*(S6), 1232–1263. http://dx.doi.org/10.1111/cogs.12384

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the selective reanalysis hypothesis. *Journal of Memory and Language, 59,* 266–293. http://dx.doi.org/10.1016/j.jml.2008.06.002

Moll, K., & Landerl, K. (2010). *SRST-II: Lese- und Rechtschreibtest: Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT)* [SRST-II: Reading and spelling test: Further development of the Salzburg Reading and Spelling Test (SRST)]. Gottingen, Germany: Hogrefe.

Pagán, A., Blythe, H. I., & Liversedge, S. P. (2016). Parafoveal preprocessing of word initial trigrams during reading in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 411–432. http://dx.doi.org/10.1037/xlm0000175

R Development Core Team. (2015). R: A language and environment for statistical computing [Computer software]. Retrieved from www.R-project.org/

Rayner, K. (1977). Visual attention in reading: Eye movements reflect cognitive processes. *Memory & Cognition, 5,* 443–448. http://dx.doi.org/10.3758/BF03197383

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

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C., Jr. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes, 4,* SI21–SI49. http://dx.doi.org/10.1080/01690968908406362

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1290–1301. http://dx.doi.org/10.1037/0278-7393.30.6.1290

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

Schroeder, S., Hyönä, J., & Liversedge, S. P. (2015). Developmental eye-tracking research in reading: Introduction to the special issue. *Journal of Cognitive Psychology, 27,* 500–510. http://dx.doi.org/10.1080/20445911.2015.1046877

Schroeder, S., Würzner, K. M., Heister, J., Geyken, A., & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior Research Methods, 47,* 1085–1094. http://dx.doi.org/10.3758/s13428-014-0528-1

Stine-Morrow, E. A., Shake, M. C., Miles, J. R., Lee, K., Gao, X., & McConkie, G. (2010). Pay now or pay later: Aging and the role of boundary salience in self-regulation of conceptual integration in sentence processing. *Psychology and Aging, 25,* 168–176. http://dx.doi.org/10.1037/a0018127

Stine-Morrow, E. A., Soederberg Miller, L. M., Gagne, D. D., & Hertzog, C. (2008). Self-regulated reading in adulthood. *Psychology and Aging, 23,* 131–153. http://dx.doi.org/10.1037/0882-7974.23.1.131

Stine-Morrow, E. A., Soederberg Miller, L. M., & Leno, R. (2001). Patterns of on-line resource allocation to narrative text by younger and older readers. *Aging, Neuropsychology and Cognition, 8,* 36–53. http://dx.doi.org/10.1076/anec.8.1.36.848

Stracuzzi, D., & Kinsey, J. (2006). EyeTrack (Version 0.7.10h) [Computer software]. Retrieved from www.psych.umass.edu/eyelab/software/

Tiffin-Richards, S. P., & Schroeder, S. (2015). Children's and adults' parafoveal processes in German: Phonological and orthographic effects. *Journal of Cognitive Psychology, 27,* 531–548. http://dx.doi.org/10.1080/20445911.2014.999076

van der Schoot, M., Vasbinder, A. L., Horsley, T. M., & Van Lieshout, E. C. (2008). The role of two reading strategies in text comprehension: An eye fixation study in primary school children. *Journal of Research in Reading, 31,* 203–223. http://dx.doi.org/10.1111/j.1467-9817.2007.00354.x

van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.

van Silfhout, G., Evers-Vermeul, J., Mak, W. M., & Sanders, T. J. (2014). Connectives and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology, 106,* 1036–1048. http://dx.doi.org/10.1037/a0036293

Warren, T., White, S. J., & Reichle, E. D. (2009). Investigating the causes of wrap-up effects: Evidence from eye movements and E–Z Reader. *Cognition, 111,* 132–137.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science, 6,* 292–297. http://dx.doi.org/10.1111/j.1467-9280.1995.tb00513.x

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185. http://dx.doi.org/10.1037/0033-2909.123.2.162

Received March 1, 2017
Revision received September 27, 2017
Accepted September 27, 2017 ■