

Title:

Beginning readers might benefit from digital texts presented in a sentence-by-sentence fashion. But why?

Highlights

- In a segmented layout, texts are presented chunk by chunk.
- Pupils in Grades 2 and 3 (7-9 years old) benefit from a segmented layout.
- Segmented texts probe higher-order comprehension by inducing more accurate reading.
- Reading skills and text genre influence the efficacy of the layout of a text.
- Segmented texts are particularly useful in the early stages of reading acquisition.

Abstract

The current digital era offers many possibilities to modify the layout of a text to optimize reading and improve comprehension. Here, we examined the idea that the visuo-spatial properties of segmented layouts support beginning readers by reducing the demands of basic eye–movement processes. In a series of self-paced reading experiments, text comprehension and reading speed of second- and third-grade pupils (N=348) were assessed in a baseline condition (i.e., sentences continued on the same line as far as page width allowed) and three conditions with a segmented layout: (1) a discontinuous layout in which each sentence was presented on a new line of the page; (2) a reader-paced Rapid Serial Visual Presentation (RSVP) layout in which the texts were presented sentence by sentence; (3) a reader-paced RSVP layout in which the texts were presented word by word. No advantages were observed for the discontinuous layout. However, at the expense of increased reading times, robust comprehension advantages emerged for the two RSVP layouts. The observed trade-off between speed and accuracy suggests that a RSVP-based layout induces more precise reading, rather than reducing the demands on basic decoding and oculomotor control processes. These findings will be discussed in the context of individual differences in reading skills and several high-potential digital applications that aim at enhancing the abilities of (beginning) readers (e.g., Spritz, BeeLine Reader).

Keywords: reading acquisition, text layout, RSVP, text comprehension, reading speed, eye movements

1. Introduction

It would almost be an understatement to emphasize that the ability to read constitutes an important skill to master in early childhood. In addition to being an essential skill, reading is a difficult skill to acquire, demanding a precisely-timed coordination between perceptual, linguistic, and more general cognitive mechanisms (e.g., Rayner, 1998). Not surprisingly, there are many attempts in the fields of, for example, education and text design, to augment the reading processes of children. The resulting approaches can be categorized as targeting either the reader or the text. The reader-oriented approaches primarily aim at enhancing pupils' word-decoding skills and comprehension strategies (e.g., McMaster et al., 2012). The text-oriented approaches are more committed to increasing the comprehensibility of a text by optimizing its grammatical and semantic content, its local and global structure, and its layout (e.g., Land, 2009).

A widely held belief in the text-oriented approaches is that beginning readers will benefit from an 'easy' text. In easy texts, words are easily recognized and sentences are easily parsed, which should result in a better understanding of what the text is about. Accordingly, texts for novices are printed in a large font with increased spacing between letters, words, and lines (e.g., Zorzi et al., 2012). In addition, infrequent words and noncanonical sentential structures are kept to a minimum. The sentences of a text are generally short (without subordinating clauses) and presented on a new line each, to avoid line breaks in the middle of a sentence (Land, 2009). In more extreme cases of these *segmented* texts, each page contains only one or two sentences – e.g., in books for the very young. The idea behind these modifications of the layout is that they optimize the eye movements and basic decoding processes during reading, thereby potentially freeing cognitive resources for higher-order comprehension processes such as monitoring, integration, and inference generation (cf., Schneps, Thomson, Sonnert, et al., 2013; Schneps, Thomson, Chen, Sonnert, & Pomplun, 2013).

Although this hypothesis is appealing, and a driving force behind several high-potential digital reading applications (e.g., Span Limiting Tactile Reinforcement, Spritz, BeeLine Reader, WebClipRead), some of its key implications have not been examined. Moreover, the empirical evidence that is available is not always in line with the predictions of the hypothesis (Land, 2009; Van Silfhout, 2014). The present study addresses these issues by examining how the layout of digital texts affects the comprehension performance of beginning readers. More specifically, in four self-paced reading experiments we studied whether 7- to 9-year-old pupils in the Dutch primary school system (Grades 2 and 3) either benefit or experience drawbacks from a segmented presentation mode of texts.

1.1. Continuous and discontinuous text layouts

A common instance of a segmented layout is when the text is presented in its entirety, yet with each sentence starting on a new line of the same page. This discontinuous layout is most prevalent in, but not limited to, texts for beginning readers. A corpus study, for example, showed that discontinuous texts are also present regularly in the curriculum of high school students (Land, 2009; Land, Sanders, & Van den Bergh, 2008). It is not always clear why publishers opt for this presentation mode, even for more experienced readers, but anecdotal evidence suggests that publishers assume that because the texts look easy, they will also be easier to read, with higher learning outcomes as a result (Land, 2009; Van Silfhout, 2014; Van Silfhout, Evers-Vermeul, Mak, & Sanders, 2014; Van Silfhout, Evers-Vermeul, & Sanders, 2014).

This justification for favoring a discontinuous over a continuous layout is overly simplistic. On the one hand, indeed, there are several good reasons to postulate that readers will benefit from a discontinuous layout. It may function as an aid to segment the texts into separate sentences and, moreover, it avoids that clausal units are interrupted by a line break, thereby limiting parsing problems for beginning readers (LeVasseur, Macaruso, Palumba & Shankweiler, 2006). In addition to simplifying these basic reading procedures, discontinuous texts may also prompt higher-order integration processes. A well-known phenomenon in reading research is that people slow down at the end of a sentence to ensure that before moving on, all within-sentence comprehension problems are settled, and the information of the sentence can be integrated with prior information of the text (Just & Carpenter, 1980). An advantage of discontinuous texts is that the line and sentence endings coincide and together present a prominent cue to the reader that sentence wrap-up and integration processes should be initiated.

On the other hand, there are also several good reasons to postulate that readers will experience drawbacks from a discontinuous layout. First, texts in which each sentence is presented on a new line tend to cover more lines than the same texts presented in a continuous fashion. Consequently, the reader must plan and execute more return sweeps in a discontinuous text. During these long saccades, readers move their eye gaze from the end of one line to the beginning of the next. This poses a challenge for inexperienced readers because corrective eye movements are frequently required to locate the optimal starting position after a return sweep (cf., Just & Carpenter, 1980; Rayner, Schotter, Masson, Potter, & Treiman, 2016). To reduce the cognitive load of oculomotor control processes for beginning readers, return sweeps should thus be kept to a minimum. Second, in a discontinuous layout, readers may perceive the text as a list of isolated events instead of approaching it as a meaningful integrated structure. This undesirable side effect of discontinuous texts is strengthened by the tendency of publishers to keep the sentences short by omitting connectives and other linguistic coherence devices. As a result, readers will not be encouraged to construct a highly connected mental representation of discontinuous texts (Van Silfhout, Evers-Vermeul, & Sanders, 2014).

In a series of experiments with 13- and 14-year-old students, Van Silfhout and her colleagues reported no facilitative influence of a discontinuous layout on reading. In fact, text comprehension was impeded in discontinuous texts for some readers, and discontinuous texts slowed down students' reading processes (Van Silfhout, 2014; Van Silfhout, Evers-Vermeul, Mak, et al., 2014; Van Silfhout, Evers-Vermeul, & Sanders, 2014). For beginning readers there are, to our knowledge, no published studies on the matter, but some preliminary results indicate that a discontinuous layout is not felicitous for 7-year-old pupils either (Evers-Vermeul & Land, 2011).¹ The results of this study, however, were inconclusive as only comprehension measures and no complementing processing measures (e.g., reading times) were reported. As a result, there are several ways to explain the absence of a comprehension difference between the two text formats. One possibility is that beginning readers benefit from a discontinuous layout, yet the advantages for comprehension are masked by increased processing efforts of readers in the continuous layout (i.e., reflecting a speed-accuracy trade-off). Another possibility is that the advantages and disadvantages of the two text formats cancel each other out, such that the overall impact on processing and comprehension is negligible. A third possibility is that the impact of each of the advantages and disadvantages is trivial in itself. After all, transforming a continuous text into a discontinuous text, or vice versa, does not alter the reading process in any fundamental way. In both layouts, the reader moves his or her eye gaze from left to right and from top to bottom to encode the visual information on a page. Moreover, in both formats readers are free to diverge from these customary reading directions by making regressive eye movements or looking ahead into the text. Hence, this raises the question of whether more extreme versions of a segmented presentation mode will have a greater impact on reading comprehension and its underlying processes than does presenting each sentence on a new line.

1.2. Rapid Serial Visual Presentation

In the current study, the label 'segmented' captures a broad spectrum of alternate approaches to present written texts, ranging from the discontinuous texts as discussed in section 1.1 to more severely segmented Rapid Serial Visual Presentation (RSVP) approaches. In RSVP, the words, sentential units or full sentences of a text are displayed sequentially on a screen, often for a predetermined, limited amount of time. There has been a long tradition of research examining the influence of RSVP on readability and text comprehension, primarily to design and evaluate new ways to present texts on the small displays of mobile phones, pagers, and more recently, smartphones and smartwatches (e.g., Benedetto, Carbone, Pedrotti, & Fevre, 2015).

¹ Although references to unpublished work should be avoided, we decided to cite the work of Evers-Vermeul and Land because their study inspired us to conduct the series of experiments presented here.

In addition to the practical advantages of RSVP, more fundamental reading advantages are attributed to the methodology (e.g., Castelhana & Muter, 2001; Lemarié, Eyrolle, & Cellier, 2008). First, RSVP may reduce the cognitive load of oculomotor control processes by decreasing the number of saccades and eliminating the need for complex return sweeps. This influence of RSVP is most noticeable when the texts are presented word by word, but similar effects arise when larger segments are presented. For example, when the segments of RSVP are full sentences, return sweeps are less demanding than in traditional texts because readers only shift their eye gaze in a horizontal plane, not in a vertical plane (cf. Van Renswoude, Johnson, Raijmakers, & Visser, 2016). Second, RSVP may reduce the negative influence of visual crowding, which is the deleterious effect of clutter on object recognition (Whitney & Levi, 2011) and known to affect reading speed (Pelli & Tillman, 2008). Third, since many popular books on speed reading strongly advocate against making regressive eye movements, a more controversial advantage would be that RSVP reduces or fully eliminates the opportunity to reread prior sections of a text (for a critical assessment of this advantage see Rayner et al., 2016). Fourth, the fact that readers cannot reread prior passages, or can only do so in a very limited way, may impact higher order integration processes. For example, Koornneef and Van Berkum (2006) proposed that readers adapt to a chunk-by-chunk presentation mode by resorting to a more incremental strategy where readers immediately integrate the information afforded by each chunk with the information of prior text (cf. Chung-Fat-Yim, Peterson, & Mar, 2017). Because incrementally updating the mental representation of a text constitutes an important aspect of proficient reading (Rayner & Clifton, 2009), readers may benefit from RSVP as it encourages them to update their mental model more frequently than they would do otherwise.

However, many of the potential advantages of RSVP are counterbalanced by its drawbacks, especially when the texts are presented word by word. First, contrary to popular belief, eye movements may not be resource consuming at all (Rayner et al., 2016) and the suppression of eye movements may even increase cognitive load and visual fatigue in single-word RSVP methods (Benedetto et al., 2015). Second, a large body of research has shown that in traditional reading situations, parafoveal preview allows a reader to use information from more than just the currently fixated word, giving the reader a ‘head start’ (Rayner et al., 2016). Single-word RSVP methods deny the reader to take advantage of this preview effect. Third, regressive eye movements are initiated to repair a failure in comprehension and protect readers against moving on without correcting their misinterpretations (Schotter, Tran, & Rayner, 2014). So, rather than causing problems, regressive eye movements are the solution to a problem (Rayner et al., 2016). Fourth, due to a lack of peripheral cues in RSVP, readers often report the feeling of ‘being lost’ (Castelhana & Muter, 2001). In fact, readers may get lost to such an extent that a RSVP text is perceived as an amorphous stream of words, rather than a meaningful integrated textual structure.

Most studies that examined the influence of RSVP-based methods on readability and text comprehension revealed a superiority of traditional reading over RSVP (for recent overviews see Benedetto et al., 2015; Lemarié et al., 2008; Rayner et al., 2016). As pointed out by Ricciardi and Di Nocera (2017), presentation speed appears to be the most important aspect to take into consideration when comparing traditional reading with RSVP, because significant reductions of reading comprehension and retention occur when reading rates exceed a critical threshold. In addition, presenting a pause at the end of each sentence of a RSVP text is crucial to approximate the comprehension scores of traditional texts, in particular for readers with a lower working memory capacity (Busler & Lazarte, 2017).

However, a few early studies revealed beneficial effects of RSVP-based methods (for an overview see Young, 1984). In addition, a recent study on adult readers showed a small increase in comprehension performance for texts that were presented sentence by sentence in a reader-controlled manner (Chung-Fat-Yim et al., 2017). Furthermore, Chen (1986) showed that readers with a low working memory span benefit from RSVP reading. Low working memory span readers performed significantly worse than readers with a high working memory span in a traditional reading condition, yet they were no obvious performance differences in the RSVP condition. According to Chen (1986), this suggests that RSVP-based techniques could be useful for improving the reading abilities or strategies of less efficient readers (cf., Busler & Lazarte, 2017). Other scholars also pointed out that RSVP-based techniques could be valuable for specific populations of readers, such as novices, the visually impaired, older adults, and dyslexics (see Castelhamo & Muter, 2001; Lemarié et al., 2008, and the references therein). These proposals, however, have never been studied in a systematic way.

1.3. The present study

The discussion above revealed that a segmented text (in its various formats) may have both positive and negative effects on basic and higher-order reading processes. It also became clear that more experienced readers do not benefit from a segmented presentation mode. In fact, for these readers the drawbacks seem to outweigh the benefits. It is not self-evident that same will hold for beginning readers because they are still acquiring and optimizing their reading skills. Since there are hardly any published studies that explored the potential value of segmented texts for novices, the first main aim of our study was straightforward. We examined whether young, beginning readers either benefit or experience drawbacks from segmented texts. We did so for mildly to more severely segmented texts, to obtain a comprehensive picture of the efficacy of adjusting the layout for beginning readers. A second aim was to explore why beginning readers may benefit from a specific layout as the visuo-spatial features of a text can affect reading in several ways, ranging from reducing the cognitive load of basic decoding and eye-movement processes to probing higher

order integration processes. A third aim was to gain more insight into whether and how individual differences in reading skills constrain the efficacy of the layout of a text.

These three aims were addressed in four self-paced reading experiments with 7- to 9-year-old Dutch pupils. In each experiment, participants read several narrative and expository texts. The reading times for the texts were recorded and each text was followed by a series of comprehension questions. In addition, standardized test scores on decoding and comprehension skills were retrieved from the schools to explore the influence of readers' individual differences. In Experiments 1-3 we compared a traditional continuous layout to three versions of a segmented layout: a discontinuous layout in which each sentence was presented on a new line of the page (Experiment 1); a RSVP-based sentence-by-sentence presentation mode in which the reader controlled when the next sentence of a text appeared on the screen by pressing a button (Experiment 2); a RSVP-based word-by-word presentation mode in which the reader controlled when the next word of a text appeared on the screen by pressing a button (Experiment 3). Experiment 4 was conducted to compare the two RSVP-based presentation modes of Experiments 2 and 3 in a single design.

2. Experiment 1: Continuous texts vs. discontinuous texts

2.1. Materials and methods

2.1.1. Participants

Participants were 81 pupils (43 girls; mean age 8.2 years; range 6.9-9.5) in Grade 2 (41 children) and Grade 3 from 21 primary schools in the Netherlands. In all experiments reported in the present study, the children had no diagnosed behavioral and/or attentional problems, and normal or corrected-to-normal vision. The parents or guardians signed a letter of active consent before testing. The children received an eraser after testing.

2.1.2. Texts and comprehension questions

Six age-appropriate texts were designed for the study, including two practice texts. The four critical texts consisted of two expository texts (one about the social structure of a community of lions and one about the human skeleton) and two narrative texts (one about children who play hide-and-seek at school and one about siblings who encounter a problem with their sister's tablet). The texts were pre-tested in a pilot study. The texts consisted of 19 sentences each and the average length was 124 words (range: 117-132 words).

To assess text comprehension, six questions of different types were posed after each text (i.e., questions tapping literal information, text-based questions requiring a text-connecting inference, and knowledge-based questions requiring a 'gap-filling' inference; see Cain & Oakhill, 1999). The answers of the children were binary scored as correct or incorrect.

2.1.3. *Decoding and comprehension proficiencies*

Two standardized, widely-used Dutch tests were administered to assess the reading skills of the pupils. The test for decoding proficiency (Three Minutes Test, developed by CITO) consisted of three lists of words of increasing difficulty. For each list, the children read aloud as many words as possible in one minute with an emphasis on both speed and accuracy. The test for reading comprehension (Reading Comprehension Test, developed by CITO) contained a range of different types of items tapping comprehension (e.g., shuffled stories, fill-in assignments, closed- and open-ended questions). In both tests, children received overall standardized scores ranging from 1 (very good) to 5 (poor). The tests were administered by the schools. The parents or guardians signed a separate letter of active consent for using the test scores in the current study.

2.1.4. *Design and procedure*

The freely accessible server (password protected) Ibex Farm (Drummond, 2013) and its supplementary software were used to run the reading experiment on a laptop at the schools of the participants. The experiment ran in the full-screen modus of an internet browser (i.e., Google Chrome, Mozilla Firefox or Apple Safari) and consisted of two main blocks. Both blocks started with oral instructions and a practice text to familiarize the participants with the procedure of each block. The practice phase of a block was followed by a testing phase in which the children read two texts for comprehension (one narrative text, one expository text). In one of the blocks, the texts were presented in their entirety, and sentences continued on the same line as far as page width allowed (*continuous* presentation mode, see Figure 1A). In the other block, the texts were presented in their entirety as well, yet line breaks in the middle of a sentence were removed as each sentence started on a new line (*discontinuous* presentation mode, see Figure 1B). The texts were presented in a white (initially empty) text box on a blue background, using a sans-serif font. The children were instructed to press the space bar to make a text appear on the computer screen. At the moment the children finished reading the text, they pressed the space bar again to progress to the comprehension questions. The elapsed time between space bar presses was recorded to obtain the total reading time for a text. After each text, six comprehension questions appeared on screen one by one. This section of the experiment was not self-paced. The test leader read out aloud the question and recorded the answer of the child by typing its content in a response text box on the screen (see Figure 2 for a schematic overview of a single trial). The ordering of the two experimental blocks and the four critical texts was rotated across four counterbalanced lists. Participants were randomly assigned to one of those lists.

Figure 1. Examples of the different presentation modes in Experiments 1-4. (Fig. A) Screenshot of the *continuous* (control) condition in Experiments 1-3. (Fig. B) Screenshot of the *discontinuous* condition in Experiment 1. (Fig. C) Screenshot of the *sentence-by-sentence* condition in Experiments 2 and 4. (Fig. D) Screenshot of the (*stationary*) *word-by-word* condition in Experiments 3 and 4. (Fig. E) Screenshot of the *moving-window word-by-word* condition in Experiment 4.

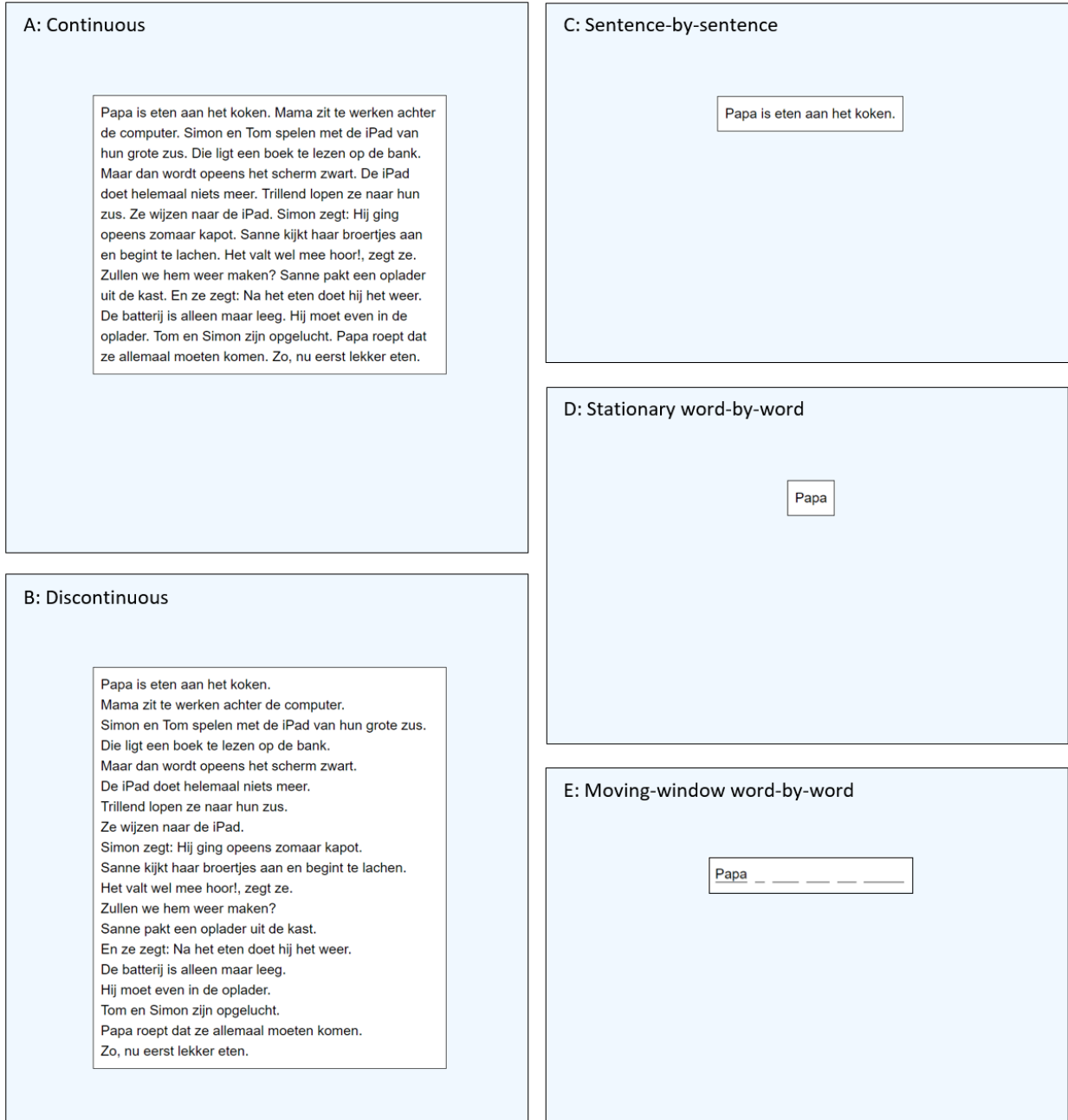
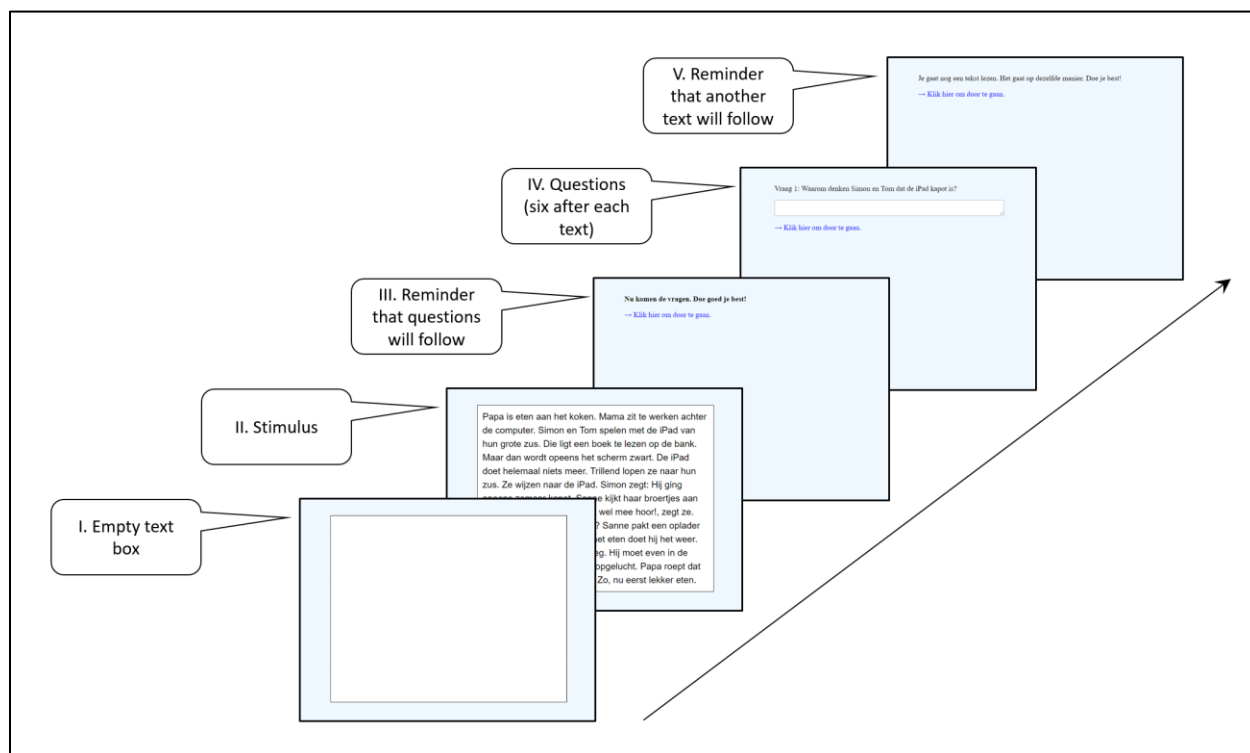


Figure 2. Time course of a single trial. A trial started with an empty text box (I). When ready, the participant pressed the space bar and the text appeared (II). When the participant was finished reading the text, he or she pressed the space bar again (note that in the sentence-by-sentence and word-by-word conditions of Experiments 2-4 the participant pressed the space bar repeatedly to progress through a text). After a reminder (III), six comprehension questions appeared one by one on the screen (IV). A trial ended with a reminder that another text (i.e., a trial of the same type) was about to follow (V) – or new instructions were provided when the participant entered the next block of the experiment.



2.2. Results

Tables 1 and 2 report the results for the comprehension questions and reading times respectively. The reading times reflect the average reading time (in milliseconds) of a word in a text. For the analyses, the reading times were log-transformed to correct for right skewness. Mixed-effects logistic regression models were fitted for the comprehension questions and mixed-effects linear regression models were fitted for the reading times. The models were fitted with the statistical software R (version 3.3.3; R Core Team, 2017) using the package LME4 (version 1.1-12; Bates, Mächler, Bolker, & Walker, 2015).

For some participants we were unable to obtain the standardized test scores for decoding and reading comprehension skills (either the parents or guardians did not give permission to use the results of these tests, or the school did not administer the tests as part of their curriculum). Consequently, the analyses consisted of two parts. In the first part we included all children that participated in the experiment and

examined the effects of PRESENTATION MODE, GRADE and TEXT GENRE. In the second part we included DECODING PROFICIENCY and COMPREHENSION PROFICIENCY as continuous predictors to explore the potential modulating effect of more basic and higher-level reading skills (note that these latter series of analyses were conducted on subsets of the total sample of participants).

Table 1. Mean accuracy scores (probability correct) for the comprehension questions in Experiments 1-3 as a function of PRESENTATION MODE, GRADE and TEXT GENRE.²

Experiment	Presentation Mode	Grade 2		Grade 3	
		Expository	Narrative	Expository	Narrative
1	Continuous	.42	.61	.57	.70
	Discontinuous	.47	.64	.53	.74
2	Continuous	.46	.63	.55	.69
	Sentence-by-sentence	.53	.66	.64	.72
3	Continuous	.45	.64	.62	.72
	Word-by-word	.50	.70	.64	.79

Table 2. Mean reading times (in milliseconds per word) and standard deviations (SD) for the texts presented in Experiments 1-3 as a function of PRESENTATION MODE, GRADE and TEXT GENRE.

Experiment	Presentation Mode	Grade 2				Grade 3			
		Expository		Narrative		Expository		Narrative	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	Continuous	683	239	628	199	538	202	510	186
	Discontinuous	682	235	653	224	524	207	490	164
2	Continuous	684	283	610	229	503	170	466	149
	Sentence-by-sentence	695	235	660	211	571	173	547	163
3	Continuous	675	300	631	235	464	152	424	146
	Word-by-word	1041	324	966	271	847	281	808	218

² Standard deviations (SD) do not apply to binary data and will only be reported for the continuous dependent variables in Tables 2 and 4.

2.2.1. *Effects of PRESENTATION MODE, GRADE and TEXT GENRE*

Mixed-effects model selection was carried out in the following way. First, a model was fitted that included the fixed factors PRESENTATION MODE (two levels: continuous vs. discontinuous), TEXT GENRE (two levels: narrative vs. expository) and GRADE (two levels: Grade 2 vs. Grade 3) and the full interactional terms for these factors. Participants and items were included as crossed random effects (Baayen, Davidson, & Bates, 2008). Wald chi-square testing (TYPE II) as implemented in the R-package CAR (version 2.1-4; Fox & Weisberg, 2011) was applied to select the most parsimonious structure of fixed effects by removing non-significant ($p > .05$) predictors. Subsequently, we included the maximal random-slope structures for participants and items that resulted in a converging model (for discussion cf. Barr, Levy, Scheepers, & Tily, 2013; Bates, Kliegl, Vasishth, & Baayen, 2015).³ For these final models, the relevant fixed-effects estimates and the associated t-values (for the continuous dependent variables) and z-values (for the categorical dependent variables) will be reported. Statistical significance at approximately the .05 level is indicated by z- and t-values of ≤ -1.96 or ≥ 1.96 . To obtain fixed-effects estimates and the associated statistics for the relevant simple effects of an interaction, the reference category of the models was adjusted and analogous models with the same structure of fixed effects and random slopes were fitted.

2.2.1.1. *Comprehension questions*

The Wald chi-square tests revealed main effects of GRADE ($\chi^2(1)=5.45$, $p < .05$) and TEXT GENRE ($\chi^2(1)=4.59$, $p < .05$).⁴ The children in Grade 3 performed better on the comprehension questions than the children in Grade 2 ($b=0.67$, $SE=0.27$, $z=2.47$). Furthermore, the performance on questions for narrative texts was better than the performance on questions for expository texts ($b=1.08$, $SE=0.52$, $z=2.07$).

2.2.1.2. *Reading times*

We observed main effects of GRADE ($\chi^2(1)=11.9$, $p < .001$) and TEXT GENRE ($\chi^2(1)=3.95$, $p < .05$).⁵ The children in Grade 3 were reading at a higher pace than the children in Grade 2 ($b=-0.26$, $SE=0.075$, $t=-3.45$) and narrative texts were read more quickly than expository texts ($b=-0.056$, $SE=0.028$, $t=-1.96$).

³ We opted for this two-step approach to simplify the procedure for fitting the random slopes, yet we still avoid inflated TYPE I error rates, as is presumably the case when only random intercepts are included (Barr et al., 2013).

⁴ Final model: Score $\sim 1 + \text{GRADE} + \text{TEXT GENRE} + (1 + \text{GRADE} + \text{TEXT GENRE} \mid \text{Participants}) + (1 + \text{GRADE} + \text{TEXT GENRE} \mid \text{Questions})$.

⁵ Final model: Log Reading Time $\sim 1 + \text{GRADE} + \text{TEXT GENRE} + (1 + \text{GRADE} + \text{TEXT GENRE} \mid \text{Participants}) + (1 \mid \text{Texts})$.

2.2.2. *Exploring interactions of DECODING and COMPREHENSION PROFICIENCIES*

The test scores on reading proficiency provided by the schools are standardized scores adjusted for grade level. As a result, the scores are not comparable across grades, and separate analyses for Grades 2 and 3 were carried out. Furthermore, because missing values for the test scores on decoding and comprehension skills were not equally distributed, we also carried out separate analyses for these two aspects of reading proficiency. In all, four separate series of analyses were conducted for the two dependent measures (i.e., accuracy comprehension questions and reading times).

Model selection was approached in a similar way as described in section 2.2.1. First, a model was fitted that included a continuous predictor of reading proficiency (either DECODING PROFICIENCY or COMPREHENSION PROFICIENCY), the factors PRESENTATION MODE and TEXT GENRE, and their full interactional terms. Participants and items were included as crossed random effects. Wald chi-square testing was applied to select the most parsimonious structure of fixed effects, followed by an inclusion of the maximal converging random-slope structures for participants and items. The continuous predictor was centered on the mean. In case of a significant modulating effect of reading proficiency, equivalent models with the relevant predictor centered to the highest (i.e. 1) and lowest (i.e. 5) proficiency scores were used to further interpret the interaction effect. We will only discuss (two- and three-way) interactions including the effects of DECODING/COMPREHENSION PROFICIENCY and PRESENTATION MODE.

2.2.2.1. *Comprehension questions*

Grade 2. There was no modulating effect of the predictor DECODING PROFICIENCY, but COMPREHENSION PROFICIENCY interacted with PRESENTATION MODE ($\chi^2(1)=5.73$, $p<.05$).⁶ Further inspection of this interaction showed that high-comprehending children performed better in the discontinuous than in the continuous condition ($b=0.63$, $SE=0.30$, $z=2.13$). This advantage diminished as a function of COMPREHENSION PROFICIENCY, to such an extent that low-comprehending children showed no comprehension advantage (numerically even a disadvantage) for the discontinuous condition ($b=-0.72$, $SE=0.44$, $z=-1.62$).

Grade 3. The analyses for Grade 3 revealed no modulating effects of DECODING PROFICIENCY and COMPREHENSION PROFICIENCY.

⁶ Final model: Score ~ 1 + PRESENTATION MODE X COMPREHENSION PROFICIENCY + TEXT GENRE + (1 + PRESENTATION MODE | Participants) + (1 | Questions).

2.2.2.2. Reading times

Grade 2. There was no modulating effect of the predictor DECODING PROFICIENCY, but COMPREHENSION PROFICIENCY interacted with PRESENTATION MODE ($\chi^2(1)=7.56, p<.01$).⁷ Further inspection of this interaction revealed that low-comprehending readers displayed longer reading times in the discontinuous condition than in the continuous condition ($b=0.11, SE=0.051, t=2.07$). High-comprehending children, on the other hand, showed no reading time difference between the continuous and discontinuous conditions (numerically they had the tendency to read the stories in the discontinuous condition more quickly: $b=-0.060, SE=0.034, t=-1.76$).

Grade 3. The analyses for Grade 3 revealed no modulating effects of DECODING PROFICIENCY and COMPREHENSION PROFICIENCY.

2.3. Discussion

In line with a number of empirical studies, Experiment 1 revealed no comprehension advantage for texts in which each sentence was presented on a new line of the page (Van Silfhout, 2014; Van Silfhout, Evers-Vermeul, Mak, et al., 2014; Van Silfhout, Evers-Vermeul, & Sanders, 2014). In addition, the results extended prior findings in two ways. First, the only study on the topic that tested beginning readers (Evers-Vermeul & Land, 2011) did not rule out that a comprehension advantage for discontinuous texts was obscured due to children's increased processing efforts (as indicated by increased reading times) when reading continuous texts. Experiment 1 showed that this explanation – which presupposes the presence of a speed-accuracy trade-off – is unwarranted as we observed no overall reading time differences between continuous and discontinuous texts. Second, our exploratory analyses on the influence of individual differences revealed an interesting pattern. High-comprehending second-grade readers performed better on the comprehension questions posed after a discontinuous text in comparison to the questions posed after a continuous text. Moreover, this comprehension advantage could not be attributed to more intensive reading in discontinuous texts – in fact, high-comprehending second-grade readers displayed a nonsignificant tendency towards faster processing times for discontinuous texts than for continuous texts. In contrast, low-comprehending second-grade readers slowed down their reading pace when confronted with discontinuous texts, perhaps because more return sweeps were required in these texts. In addition, these increased processing costs were not accompanied by higher comprehension scores. Hence, whereas proficient beginning readers benefit from discontinuous texts, struggling beginning readers do not.

⁷ Final model: $\text{Log Reading Time} \sim 1 + \text{PRESENTATION MODE} \times \text{COMPREHENSION PROFICIENCY} + \text{TEXT GENRE} + (1 + \text{PRESENTATION MODE} + \text{TEXT GENRE} | \text{Participants}) + (1 + \text{PRESENTATION MODE} \times \text{COMPREHENSION PROFICIENCY} + \text{TEXT GENRE} | \text{Texts})$.

3. Experiment 2: Continuous texts vs. sentence-by-sentence texts

3.1. Materials and methods

3.1.1. Participants

Participants were 88 pupils (47 girls; mean age 8.2 years; range 7.0-9.4) in Grades 2 (46 children) and 3 from 21 primary schools in the Netherlands. None of them participated in Experiment 1.

3.1.2. Materials

The stimuli and standardized tests for decoding and comprehension proficiencies were identical to those of Experiment 1.

3.1.3. Design and procedure

Experiment 2 also consisted of two experimental blocks. One block was identical to the *continuous* condition of Experiment 1 (see Figure 1A). The main (and only) difference with Experiment 1 was that the experimental block in which each sentence was presented on a new line, was replaced by a condition in which each sentence of the texts was presented separately (*sentence-by-sentence* condition, see Figure 1C). In this condition a trial also started with a white (yet smaller) empty text box in the middle of the screen. After the child pressed the space bar, the first sentence of a text appeared in the box. After pressing the space bar again, the first sentence of a text was replaced by the second sentence of that text. By repeatedly pressing the space bar the child read all the sentences of a text in a sentence-by-sentence manner. After pressing the space bar, it was not possible to go back to sentences presented earlier in the trial. The elapsed time between space bar presses was recorded to obtain reading times for the sentences of a text.

3.2. Results

Tables 1 and 2 report the results for the comprehension questions and reading times respectively. The procedure for the analyses was identical to Experiment 1.

3.2.1. *Effects of PRESENTATION MODE, GRADE and TEXT GENRE*

3.2.1.1. *Comprehension questions*

We observed a main effect of PRESENTATION MODE ($\chi^2(1)=8.22, p<.01$).⁸ Relative to the continuous condition, children performed better in the sentence-by-sentence condition ($b=0.32, SE=0.15, z=2.18$).

3.2.1.2. *Reading times*

We observed main effects of PRESENTATION MODE ($\chi^2(1)=57.7, p<.001$), GRADE ($\chi^2(1)=9.14, p<.01$) and TEXT GENRE ($\chi^2(1)=7.43, p<.01$).⁹ Children were reading at a slower pace in the sentence-by-sentence condition than in the continuous condition ($b=0.11, SE=0.026, t=4.24$), children in Grade 3 were reading at a higher pace than children in Grade 2 ($b=-0.20, SE=0.075, t=-2.64$), and narrative texts were read more quickly than expository texts ($b=-0.045, SE=0.019, t=-2.36$).

3.2.2. *Exploring interactions of DECODING and COMPREHENSION PROFICIENCIES*

3.2.2.1. *Comprehension questions*

The analyses revealed no modulating effects of DECODING PROFICIENCY and COMPREHENSION PROFICIENCY in Grades 2 and 3.

3.2.2.2. *Reading times*

Grade 2. We observed a three-way interaction of DECODING PROFICIENCY X PRESENTATION MODE X GENRE ($\chi^2(1)=5.68, p<.05$).¹⁰ Narrative texts (see Figure 3, top left) presented in a sentence-by-sentence manner were read more slowly than narrative texts presented in a continuous manner ($b=0.11, SE=0.034, t=3.14$). There was no DECODING PROFICIENCY X PRESENTATION MODE interaction for narrative texts ($b=-0.023, SE=0.021, t=-1.07$). In contrast, for expository texts (see Figure 3, top right) DECODING PROFICIENCY interacted with PRESENTATION MODE ($b=-0.82, SE=0.021, t=-3.89$). More specifically, proficient decoders displayed longer reading times for the sentence-by-sentence condition than for the

⁸ Final model: Score ~ 1 + PRESENTATION MODE + (1 + PRESENTATION MODE | Participants) + (1 + PRESENTATION MODE | Questions).

⁹ In addition, the Wald chi-square tests revealed a PRESENTATION MODE X GRADE interaction ($\chi^2(1)=7.79, p<.01$), but this interaction fell short of significance after including the random slopes ($\chi^2(1)=3.63, p=.057$). Final model: Log Reading Time ~ 1 + PRESENTATION MODE + GRADE + TEXT GENRE + (1 + PRESENTATION MODE + TEXT GENRE | Participants) + (1 + PRESENTATION MODE + GRADE + TEXT GENRE | Texts).

¹⁰ Final model: Log Reading Time ~ 1 + PRESENTATION MODE X TEXT GENRE X DECODING PROFICIENCY + (1 + PRESENTATION MODE + TEXT GENRE | Participants) + (1 + PRESENTATION MODE X TEXT GENRE + DECODING PROFICIENCY | Texts).

continuous condition ($b=0.15$, $SE=0.044$, $t=3.54$), yet struggling decoders displayed the opposite pattern (i.e., shorter reading times for the sentence-by-sentence condition, $b=-0.17$, $SE=0.065$, $t=-2.65$).

In addition, we observed a three-way interaction of COMPREHENSION PROFICIENCY X PRESENTATION MODE X GENRE ($\chi^2(1)=4.80$, $p<.05$).¹¹ As depicted in Figure 3 (bottom graphs), narrative texts in the sentence-by-sentence condition were read more slowly than narrative texts in the continuous condition ($b=0.098$, $SE=0.037$, $t=2.68$). This effect of PRESENTATION MODE did not interact with COMPREHENSION PROFICIENCY ($b=0.0065$, $SE=0.025$, $t=0.26$). For the expository texts, however, COMPREHENSION PROFICIENCY interacted with PRESENTATION MODE ($b=-0.055$, $SE=0.025$, $t=-2.24$). Whereas high-comprehending readers displayed longer reading times for the sentence-by-sentence condition than for the continuous condition ($b=0.13$, $SE=0.059$, $t=2.16$), low-comprehending readers did not display a reading time difference between these conditions ($b=-0.094$, $SE=0.069$, $t=-1.36$).

Grade 3. The analyses revealed no modulating effects of DECODING PROFICIENCY and COMPREHENSION PROFICIENCY.

3.3. Discussion

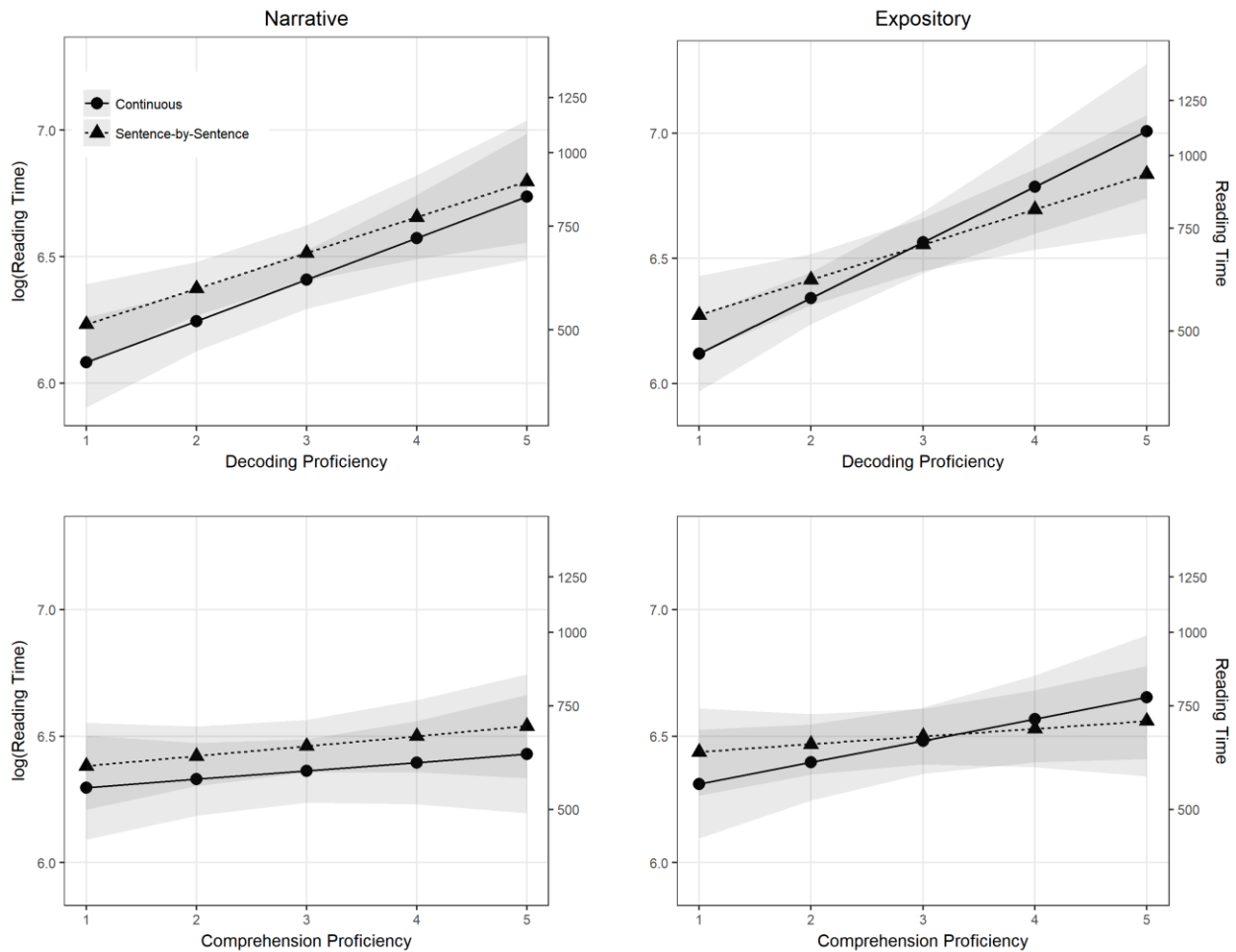
Experiment 2 showed that beginning readers have a better understanding of a text when presented in a self-paced sentence-by-sentence manner than when presented in a traditional continuous manner. In addition, the analyses of the reading times showed that sentence-by-sentence texts were processed more slowly than continuous texts. Together, these results suggest a speed-accuracy trade-off in which the higher comprehension scores for sentence-by-sentence texts should at least partly be attributed to a more laborious processing style of the readers.

The exploratory analyses on the influence of individual differences, however, sketched a more complicated picture. In the case of expository texts, struggling second-grade readers did not adhere to the above-mentioned speed-accuracy trade-off as they processed sentence-by-sentence expository texts more quickly than continuous expository texts, apparently without severely compromising their understanding of the content of the text. So, a more comprehensive conclusion would be that, in general, beginning readers benefit from a text presented in segments because this will induce a more accurate, resource-consuming processing strategy. An exception to this general rule applies to struggling second-grade pupils when they read expository texts. In that specific situation a more plausible advantage of the sentence-by-sentence

¹¹ Final model: $\text{Log Reading Time} \sim 1 + \text{PRESENTATION MODE} \times \text{TEXT GENRE} \times \text{COMPREHENSION PROFICIENCY} + (1 + \text{PRESENTATION MODE} + \text{TEXT GENRE} \mid \text{Participants}) + (1 + \text{PRESENTATION MODE} \times \text{TEXT GENRE} + \text{COMPREHENSION PROFICIENCY} \mid \text{Texts})$.

presentation method is that the processing load of basic reading processes (e.g., parsing, eye-movement control) is reduced – which in turn may have a beneficial influence on comprehension.

Figure 3. Fixed effects estimates (and their 95% confidence intervals) of the log-transformed reading times (in milliseconds per word) of second-grade pupils in Experiment 2, as a function of PRESENTATION MODE, DECODING PROFICIENCY (top), COMPREHENSION PROFICIENCY (bottom) and TEXT GENRE (narrative texts on the left, expository texts on the right). Scales of exponentiated log-values (i.e., approximating untransformed values) are provided as secondary y-axes on the right side of the graphs.¹²



¹² The fixed effects and confidence intervals were extracted and plotted with the R-packages EFFECTS (Fox, n.d.; Fox & Hong, 2009) and GGPLOT2 (Wickham, 2009).

4. Experiment 3: Continuous texts vs. word-by-word texts

4.1. Materials and methods

4.1.1. Participants

Participants were 83 pupils (45 girls; mean age 8.2 years; range 7.1-9.6) in Grades 2 (46 children) and 3 from 21 primary schools in the Netherlands. None of them participated in the previous experiments.

4.1.2. Materials

The stimuli and standardized tests for decoding and comprehension proficiency were identical to those of the previous experiments.

4.1.3. Design and procedure

Experiment 3 also consisted of two experimental blocks. Again, one of the blocks was identical to the *continuous* condition. In the other block the participants read the texts in a word-by-word manner (see Figure 1D). A trial started with a small empty text box. After the child pressed the space bar, the first word of a text appeared in this box. By pressing the space bar another time, the first word of the text was replaced by its second word. The child read all the words of a text by repeatedly pressing the space bar. It was not possible to go back to words presented earlier in a trial. The elapsed time between space bar presses was recorded to obtain reading times for the words of a text.

4.2. Results

Tables 1 and 2 report the results for the comprehension questions and reading times respectively. The procedure for the analyses was identical to the previous experiments.

4.2.1. Effects of PRESENTATION MODE, GRADE and TEXT GENRE

4.2.1.1. Comprehension questions

We observed main effects of PRESENTATION MODE ($\chi^2(1)=8.96, p<.01$), GRADE ($\chi^2(1)=10.31, p<.01$), and TEXT GENRE ($\chi^2(1)=5.69, p<.05$). Children answered more questions correctly in the word-by-word condition than in the continuous condition ($b=0.33, SE=0.11, z=3.00$), children in Grade 3 performed better on the comprehension questions than children in Grade 2 ($b=0.71, SE=0.23, z=3.13$), and the performance

on questions for narrative texts was better than the performance on questions for expository texts ($b=1.02$, $SE=0.44$, $z=2.33$).¹³

4.2.1.2. Reading times

We observed main effects of PRESENTATION MODE ($\chi^2(1)=542$, $p<.001$) and GRADE ($\chi^2(1)=14.2$, $p<.001$), and a PRESENTATION MODE X GRADE interaction ($\chi^2(1)=12.7$, $p<.001$).¹⁴ Children in Grade 3 were reading at a higher pace than children in Grade 2 in both conditions (continuous: $b=-0.35$, $SE=0.091$, $t=-3.81$; word-by-word: $b=-0.17$, $SE=0.067$, $t=-2.55$). Furthermore, children in both grades displayed longer reading times in the word-by-word condition than in the continuous condition (Grade 2: $b=0.48$, $SE=0.055$, $t=8.68$; Grade 3: $b=0.65$, $SE=0.062$, $t=10.57$). The PRESENTATION MODE X GRADE interaction indicated that, relative to the continuous condition, children in Grade 3 showed a more marked reading time increase in the word-by-word condition than the children in Grade 2 did.

4.2.2. Exploring interactions of DECODING and COMPREHENSION PROFICIENCIES

The analyses for the comprehension questions and reading times revealed no modulating effects of DECODING PROFICIENCY and COMPREHENSION PROFICIENCY in Grades 2 and 3.

4.3. Discussion

Experiment 3 mirrored the findings of Experiment 2 in showing that beginning readers obtained a better understanding of a text that was presented chunk by chunk, relative to the situation in which the text was presented in its entirety. Experiment 3 extended the findings of Experiment 2 by demonstrating that a comprehension advantage persisted when smaller chunks were presented, consisting of only a single word. In addition, the results of Experiment 3 displayed a speed-accuracy trade-off as the reading times for the word-by-word texts were considerably longer than the reading times for the continuous texts – for third-grade readers the reading times for the word-by-word texts were almost twice as long.

5. Experiment 4: Sentence-by-sentence texts vs. word-by-word texts

¹³ Final model: Score $\sim 1 + \text{PRESENTATION MODE} + \text{GRADE} + \text{TEXT GENRE} + (1 + \text{GRADE} + \text{TEXT GENRE} | \text{Participants}) + (1 + \text{TEXT GENRE} | \text{Questions})$.

¹⁴ Final model: Log Reading Time $\sim 1 + \text{PRESENTATION MODE} \times \text{GRADE} + (1 + \text{PRESENTATION MODE} + \text{GRADE} | \text{Participants}) + (1 + \text{PRESENTATION MODE} \times \text{GRADE} | \text{Texts})$.

Experiments 2 and 3 revealed that sentence-by-sentence and word-by-word texts increase reading comprehension in comparison to a continuous text. It is, however, difficult to decipher how the two segmented texts induce this increment. Moreover, it is unclear which of the two segmented presentation modes reflects the most suitable way of presenting texts to beginning readers because they were examined in two distinct experiments. To address these issues, we directly compared the influence on reading comprehension of sentence-by-sentence and word-by-word presentation modes in Experiment 4. In addition, we included a third condition in which the texts were presented in a self-paced word-by-word moving-window fashion (henceforth, we will refer to the word-by-word condition of Experiment 3 as the *stationary* word-by-word condition and we will refer to the new condition as the *moving-window* word-by-word condition) (cf. Busler & Lazarte, 2017). In this condition, each word of a sentence appeared at the same position it would occupy if the sentence was presented in its entirety. Accordingly, the children read each sentence of the text in a word-by-word manner, yet they were still required to plan and execute saccades, comparable to the situation in which the consecutive segments of the text were full sentences. In all, Experiment 4 allowed us to more carefully consider the influence of the following factors on reading speed and text comprehension: (1) visual crowding (i.e., present in the sentence-by-sentence condition, reduced in the word-by-word conditions); (2) the word-preview effect (present in the sentence-by-sentence condition, absent in the word-by-word conditions); and (3) left-to-right (i.e., horizontal) saccadic eye movements (required in the sentence-by-sentence and moving-window conditions, not required in the stationary condition).

5.1. Materials and methods

5.1.1. Participants

Participants were 96 pupils (55 girls; mean age 7.9 years; range 6.8-9.0) in Grades 2 (43 children) and 3 from 9 primary schools in the Netherlands. None of them participated in the previous experiments.

5.1.2. Materials

Experiment 4 included three self-paced reading conditions. To ensure that in every condition two texts were read by the participants, we constructed two additional critical texts (one expository text and one narrative text) with six comprehension questions each. The expository text explained what it takes to become an astronaut (155 words). The narrative text told the story of a chubby elephant, eventually meeting the love of his life (134 words). The standardized tests for decoding and comprehension proficiencies were identical to those of the previous experiments.

5.1.3. *Design and procedure*

In addition to the sentence-by-sentence (see Figure 1C) and stationary word-by-word (see Figure 1D) conditions, the experiment included a moving-window word-by-word condition (see Figure 1E). In this latter condition, a trial started with a text box that already contained horizontal lines that indicated the length and position of the words of the first sentence of a text. When the child pressed the space bar, the first word of the sentence appeared above its corresponding line in the text box. When the child pressed the space bar for a second time, the first word of the sentence disappeared, and at the same time the second word of the sentence was presented above its own corresponding line in the text box, i.e., to the right of the first word. By repeatedly pressing the space bar, the child read the first sentence of a text by shifting his or her eye gaze from word to word in the sentence. When the child finished reading the first sentence, an empty text box with horizontal lines indicating the structure of the second sentence replaced the text box of the first sentence. The second and remaining sentences of a text were read in the same word-by-word moving window fashion as described for the first sentence. It was not possible for the child to go back to words or sentences that were presented earlier in a trial. The elapsed time between space bar presses was recorded to obtain reading times for the words of a text. The ordering of the three experimental blocks and the six critical texts was rotated across six counterbalanced lists. Participants were randomly assigned to one of those lists.

5.2. Results

Tables 3 and 4 report the results for the comprehension questions and reading times respectively. Four trials (1.2% of the data) were removed due to unrealistically short average reading times per word (< 40 milliseconds). The procedure for the analyses was identical to the previous experiments.

Table 3. Mean accuracy scores (probability correct) for the comprehension questions in Experiments 4 as a function of PRESENTATION MODE, GRADE and TEXT GENRE.

Presentation Mode	Grade 2		Grade 3	
	Expository	Narrative	Expository	Narrative
Sentence-by-sentence	.50	.60	.62	.79
Stationary word-by-word	.56	.62	.66	.78
Moving word-by-word	.53	.64	.70	.76

Table 4. Mean reading times (in milliseconds per word) and standard deviations (SD) for the texts presented in Experiments 4 as a function of PRESENTATION MODE, GRADE and TEXT GENRE.

Presentation Mode	Grade 2				Grade 3			
	Expository		Narrative		Expository		Narrative	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Sentence-by-sentence	986	525	847	373	542	232	527	166
Stationary word-by-word	1127	565	1121	580	853	252	837	237
Moving word-by-word	1008	379	964	319	757	203	732	186

5.2.1. Effects of PRESENTATION MODE, GRADE and TEXT GENRE

5.2.1.1. Comprehension questions

We observed a main effect of GRADE ($\chi^2(1)=20.6, p<.001$).¹⁵ Children in Grade 3 performed better on the comprehension questions than children in Grade 2 did ($b=0.90, SE=0.20, z=4.46$).

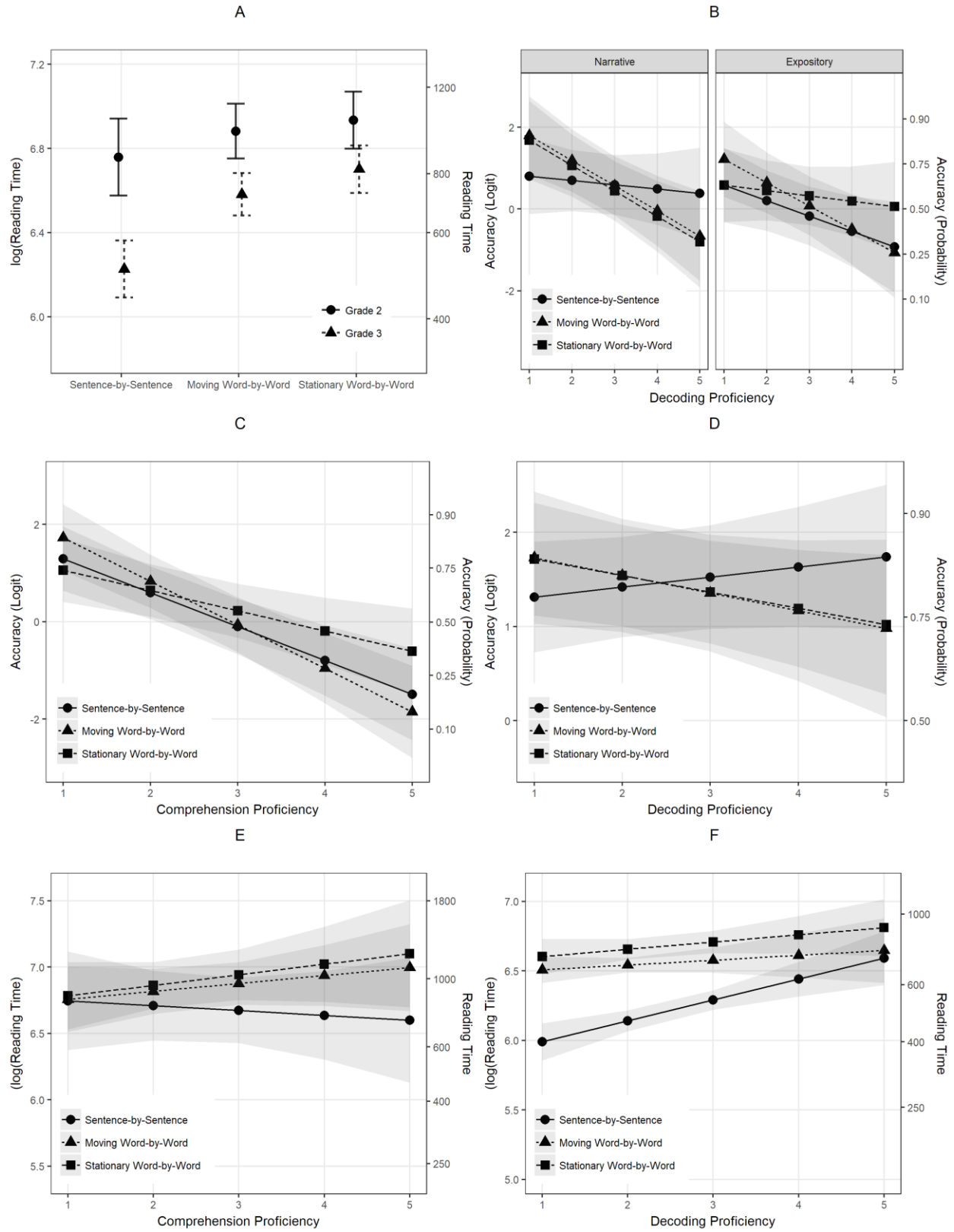
5.2.1.2. Reading times

We observed main effects of PRESENTATION MODE ($\chi^2(2)=253, p<.001$) and GRADE ($\chi^2(1)=17.6, p<.001$), and a PRESENTATION MODE X GRADE interaction ($\chi^2(2)=43.2, p<.001$).¹⁶ As illustrated in Figure 4A, pupils in Grade 3 read more quickly than pupils in Grade 2 did, in all three conditions (sentence-by-sentence: $b=-0.53, SE=0.11, t=-4.78$; stationary: $b=-0.23, SE=0.089, t=-2.63$; moving-window: $b=-30, SE=0.081, t=-3.72$). The impact of the three reading conditions, however, diverged for second- and third-grade readers. Whereas pupils in both grades read more slowly in the stationary condition than in the sentence-by-sentence condition (Grade 2: $b=0.18, SE=0.070, t=2.51$; Grade 3: $b=0.47, SE=0.052, t=9.09$), the two remaining contrasts (moving-window vs. sentence-by-sentence, moving-window vs. stationary) showed a different pattern. More specifically, pupils in Grade 3 read more slowly in the moving-window condition than in the sentence-by-sentence condition ($b=0.35, SE=0.046, t=7.70$), but pupils in Grade 2 did not display a significant increase in reading times for the moving-window condition ($b=0.12, SE=0.066, t=1.86$). Similarly, whereas readers in Grade 3 read more slowly in the stationary condition than in the moving-window condition ($b=0.12, SE=0.030, t=4.03$), pupils in Grade 2 did not display reliable reading time differences between these two conditions ($b=0.052, SE=0.038, t=1.36$).

¹⁵ Final model: Score ~ 1 + GRADE + (1 + GRADE | Participants) + (1 + GRADE | Questions).

¹⁶ Final model: Log Reading Time ~ 1 + PRESENTATION MODE X GRADE + (1 + PRESENTATION MODE | Participants) + (1 + PRESENTATION MODE X GRADE | Texts).

Figure 4. Fixed effects estimates (and their 95% confidence intervals) of the accuracy scores (probability correct) and log-transformed reading times (in milliseconds per word) of pupils in Experiment 4. The figures for the accuracy scores contain logit scales (primary y-axes in the left) and probability scales (secondary y-axes on the right), and the figures for the reading times contain log scales (primary y-axes in the left) and exponentiated log scales (secondary y-axes on the right). (*Fig. A*) Estimates of the log-transformed reading times as a function of PRESENTATION MODE and GRADE. (*Fig. B*) Estimates of the accuracy scores of second-grade pupils as a function of PRESENTATION MODE, DECODING PROFICIENCY and TEXT GENRE. (*Fig. C*) Estimates of the accuracy scores of second-grade pupils as a function of PRESENTATION MODE and COMPREHENSION PROFICIENCY. (*Fig. D*) Estimates of the accuracy scores of third-grade pupils as a function of PRESENTATION MODE and DECODING PROFICIENCY. (*Fig. E*) Estimates of the log-transformed reading times of second-grade pupils as a function of PRESENTATION MODE and COMPREHENSION PROFICIENCY. (*Fig. F*) Estimates of the log-transformed reading times of third-grade pupils as a function of PRESENTATION MODE and DECODING PROFICIENCY.



5.2.2. Exploring interactions of *DECODING* and *COMPREHENSION PROFICIENCIES*

5.2.2.1. Comprehension questions

Grade 2. We observed a reliable three-way interaction of *DECODING PROFICIENCY* X *PRESENTATION MODE* X *GENRE* ($\chi^2(2)=7.98, p<.05$).¹⁷ We discuss this interaction in the context of narrative texts first. As illustrated in Figure 4B (left graph), struggling decoders obtained a higher accuracy score for the narrative texts in the sentence-by-sentence condition in comparison to both word-by-word conditions (sentence-by-sentence vs. stationary: $b=-1.18, SE=0.53, z=-2.24$; sentence-by-sentence vs. moving-window: $b=-1.04, SE=0.52, z=-1.99$). No differences were observed between the two word-by-word conditions ($b=-0.14, SE=0.52, z=-0.28$). A different pattern emerged for proficient decoders. That is, although no difference was present between the two word-by-word conditions ($b=-0.12, SE=0.42, z=-0.27$), both conditions induced higher accuracy scores than the sentence-by-sentence condition did (sentence-by-sentence vs. stationary: $b=0.88, SE=0.41, z=2.12$; sentence-by-sentence vs. moving-window: $b=1.00, SE=0.41, z=2.41$). The results for the expository texts sketched a different picture (see Figure 4B, right graph). For struggling decoders, the highest accuracy score was observed in the stationary condition (moving-window vs. stationary: $b=1.13, SE=0.51, z=2.22$; sentence-by-sentence vs. stationary: $b=0.96, SE=0.51, z=1.90$; sentence-by-sentence vs. moving-window: $b=-0.15, SE=0.51, z=-0.29$). For proficient decoders, however, *PRESENTATION MODE* did not reliably affect the accuracy scores for expository texts (sentence-by-sentence vs. stationary: $b=0.0057, SE=0.38, z=0.015$; sentence-by-sentence vs. moving-window: $b=0.63, SE=0.39, z=1.63$; moving-window vs. stationary: $b=-0.63, SE=0.38, z=-1.64$).

In addition, *COMPREHENSION PROFICIENCY* interacted with *PRESENTATION MODE* ($\chi^2(2)=7.79, p<.05$).¹⁸ As depicted in Figure 4C, struggling comprehenders obtained the highest accuracy score while reading in a stationary manner and the lowest accuracy score while reading in a moving-window manner (moving-window vs. stationary: $b=1.24, SE=0.49, z=2.54$; sentence-by-sentence vs. stationary: $b=0.88, SE=0.49, z=1.81$; sentence-by-sentence vs. moving-window: $b=-0.36, SE=0.52, z=0.69$). Proficient comprehenders showed the opposite pattern. They obtained the lowest accuracy score in the stationary condition and the highest accuracy score in the moving-window condition (moving-window vs. stationary: $b=-0.67, SE=0.31, z=-2.13$; sentence-by-sentence vs. stationary: $b=-0.23, SE=0.30, z=-0.77$; sentence-by-sentence vs. moving-window: $b=0.43, SE=0.32, z=1.35$).

¹⁷ Due to convergence issues we could not include random slopes. Final model: Score $\sim 1 + \text{PRESENTATION MODE} \times \text{TEXT GENRE} \times \text{DECODING PROFICIENCY} + (1 | \text{Participants}) + (1 | \text{Questions})$.

¹⁸ Due to convergence issues we could not include random slopes. Final model: Score $\sim 1 + \text{COMPREHENSION PROFICIENCY} \times \text{PRESENTATION MODE} + (1 | \text{Participants}) + (1 | \text{Questions})$.

Grade 3. As illustrated in Figure 4D, DECODING PROFICIENCY interacted with PRESENTATION MODE ($\chi^2(2)=6.10, p<.05$).¹⁹ The analyses for struggling decoders revealed no significant differences between the three reading conditions, yet the general pattern was that their accuracy scores were lower in the two word-by-word conditions in comparison to the sentence-by-sentence condition (sentence-by-sentence vs. stationary: $b=-0.72, SE=0.39, z=-1.83$; sentence-by-sentence vs. moving-window: $b=-0.76, SE=0.42, z=-1.79$; moving-window vs. stationary: $b=0.033, SE=0.44, z=0.08$). Proficient decoders showed no significant differences between the three conditions either, yet for proficient decoders the two word-by-word conditions had a positive, instead of a negative impact on the accuracy scores (sentence-by-sentence vs. stationary: $b=0.40, SE=0.24, z=1.64$; sentence-by-sentence vs. moving-window: $b=0.42, SE=0.27, z=1.57$; moving-window vs. stationary: $b=-0.015, SE=0.29, z=-0.05$). We observed no reliable interactions of COMPREHENSION PROFICIENCY.

5.2.2.2. Reading times

Grade 2. We observed no reliable interactions of DECODING PROFICIENCY, but COMPREHENSION PROFICIENCY interacted with PRESENTATION MODE ($\chi^2(2)=20.2, p<.001$).²⁰ As illustrated in Figure 4E, struggling comprehenders read more quickly in the sentence-by-sentence condition, relative to both word-by-word conditions (sentence-by-sentence vs. stationary: $b=0.50, SE=0.16, t=3.18$; sentence-by-sentence vs. moving-window: $b=0.40, SE=0.15, t=2.65$; moving-window vs. stationary: $b=0.10, SE=0.076, t=1.39$). This reading time advantage for the sentence-by-sentence condition was absent for proficient comprehenders (sentence-by-sentence vs. stationary $b=0.037, SE=0.095, t=0.39$; sentence-by-sentence vs. moving-window: $b=0.010, SE=0.092, t=0.11$; moving-window vs. stationary: $b=0.023, SE=0.052, t=0.45$).

Grade 3. DECODING PROFICIENCY interacted with PRESENTATION MODE ($\chi^2(2)=36.4, p<.001$).²¹ As depicted in Figure 4F, struggling decoders read most slowly in the stationary condition (sentence-by-sentence vs. stationary: $b=0.22, SE=0.092, t=2.36$; moving-window vs. stationary: $b=0.16, SE=0.064, t=2.57$) and obtained a similar reading pace in the sentence-by-sentence and moving-window conditions ($b=0.052, SE=0.080, t=0.65$). Proficient decoders obtained their highest reading pace in the sentence-by-sentence condition (sentence-by-sentence vs. stationary $b=0.61, SE=0.066, t=9.34$; sentence-by-sentence

¹⁹ Final model: Score $\sim 1 + \text{DECODING PROFICIENCY} \times \text{PRESENTATION MODE} + \text{TEXT GENRE} + (1 + \text{PRESENTATION MODE} | \text{Participants}) + (1 | \text{Questions})$.

²⁰ Final model: Log Reading Time $\sim 1 + \text{COMPREHENSION PROFICIENCY} \times \text{PRESENTATION MODE} + (1 + \text{COMPREHENSION PROFICIENCY} + \text{PRESENTATION MODE} | \text{Participants}) + (1 + \text{COMPREHENSION PROFICIENCY} \times \text{PRESENTATION MODE} | \text{Texts})$.

²¹ Final model: Log Reading Time $\sim 1 + \text{DECODING PROFICIENCY} \times \text{PRESENTATION MODE} + (1 + \text{DECODING PROFICIENCY} + \text{PRESENTATION MODE} | \text{Participants}) + (1 + \text{DECODING PROFICIENCY} \times \text{PRESENTATION MODE} | \text{Texts})$.

vs. moving-window: $b=0.52$, $SE=0.052$, $t=9.94$; moving-window vs. stationary: $b=0.096$ $SE=0.050$, $t=1.94$). We observed no reliable interactions of COMPREHENSION PROFICIENCY.

5.3. Discussion

None of the three presentation modes revealed a clear comprehension advantage, yet texts that were presented sentence by sentence induced a significant reading time advantage over texts that were presented word by word. Hence, the sentence-by-sentence format seems to be a more efficient way of presenting texts to beginning readers. These results indicate that, in general, horizontal saccadic eye movements and visual crowding effects of adjacent words do not pose substantial problems for beginning readers. Moreover, word-preview benefits – and the opportunity to skip words – may permit readers to attain a higher reading pace without compromising their comprehension of texts.

However, the exploratory analyses on the influence of readers' decoding and comprehension proficiencies revealed a more complicated picture, especially for second-grade readers. Perhaps the most remarkable finding was that second-grade readers with lower decoding skills obtained relatively high accuracy scores for narrative texts that were presented in a sentence-by-sentence format, yet their accuracy scores for expository texts were relatively high in the stationary word-by-word format. This suggests that reducing the number of saccades may be beneficial for struggling decoders if the text poses high demands on their comprehension skills, as is presumably the case when they are confronted with an expository text – a genre younger readers are often unfamiliar with (Best, Floyd, & Mcnamara, 2008). This view on the data is strengthened by the observation that the accuracy scores of second-grade readers with lower comprehension skills were also relatively high during stationary reading, in particular when compared to their accuracy scores in the moving-window word-by-word condition.

In addition, the exploratory analyses showed that not only struggling readers benefit from word-by-word texts. Proficient second-grade decoders obtained their highest accuracy scores for narrative texts when reading in a stationary and moving-window fashion. Similarly, proficient second-grade comprehenders obtained their highest accuracy scores in the moving-window condition. However, for these readers the beneficial effects of word-by-word reading cannot be attributed to a reduction of the number of saccades. In fact, the opposite seems to be the case. Preventing proficient second-grade comprehenders from moving their eyes during reading has a detrimental effect on comprehension as they obtained their lowest comprehension scores during stationary reading. Hence, perhaps a more plausible explanation for the advantages of word-by-word reading for proficient second-grade readers is that it induces an (incremental) processing approach in which the information of each word is immediately and meticulously integrated with the information of prior text (Koorneef & Van Berkum, 2006).

Lastly, the advantages of reading without saccades (i.e., stationary reading) are most relevant in Grade 2. By the time pupils are attending Grade 3, the vast majority of pupils – including those who struggle – benefit from a text format in which a more customary reading style (i.e., including horizontal saccadic eye movements) can be attained.

6. General Discussion

Three research objectives were identified in the current study. First, we examined whether second- and third-grade readers in the Dutch educational system either benefit or experience drawbacks from texts with a segmented layout. Second, we considered the sources of the advantages and disadvantages of segmented texts by evaluating how their features impact basic reading and higher-order integration processes. Third, we studied how individual differences in reading abilities constrain the efficacy of the layout of a text.

Relative to texts with a traditional layout, no overall comprehension advantage was observed for texts in which each sentence was presented on a new line of the page (Experiment 1). Yet, a robust comprehension advantage manifested itself in two self-paced RSVP layouts in which the texts were presented one sentence or one word at a time (Experiments 2 and 3). Improved comprehension was accompanied by increased reading times in both presentation modes. This trade-off between speed and accuracy was exceptionally prominent when the texts were presented word by word. The reading time delays in texts that were presented sentence by sentence were far less extreme (Experiments 2-4). We therefore postulate that a single-sentence self-paced RSVP format constitutes a suitable way – arguably one of the most suitable ways – of presenting texts to beginning readers.

6.1. Why do beginning readers benefit from segmented texts?

As discussed in sections 1.1 and 1.2, a comprehension advantage for segmented texts may stem from two main sources. On the one hand, segmented texts may require fewer cognitive resources for basic reading (decoding, parsing) and eye-movement processes, thereby releasing cognitive resources for higher-order comprehension processes. On the other hand, segmented texts may probe higher-order comprehension by inducing a more accurate yet intensive processing approach. While both accounts have their merits, we will argue that the overall pattern of results is more compatible with the latter account.

Relating to the first account, we obtained no strong evidence for the claim that reading comprehension is compromised when clausal units are interrupted by a line break (Experiment 1) (cf., Levasseur et al., 2006; Tiffin-Richards & Schroeder, 2018). Furthermore, horizontal saccadic eye movements do not seem to pose a severe problem for beginning readers and reducing the degree of visual crowding does not positively influence reading comprehension and reading speed. At least not to the extent

that it outweighs word-preview benefits and the possibility to skip words (Experiment 4). Our results, however, do point out that beginning readers may profit from limiting or fully eliminating the possibility to look back to earlier sections of a text – as is the case in sentence-by-sentence and word-by-word presentation modes respectively. Moreover, removing the need for complex return sweeps may be beneficial for reading comprehension (Experiments 2 and 3). Hence, these findings are compatible with the claim that segmented texts positively impact comprehension by streamlining more basic aspects of reading.

However, the presence of a robust speed-accuracy trade-off in Experiments 2 and 3 lends more support for an account in which the comprehension advantage of segmented texts is attributed to more intensive processing on part of the reader. Intensive processing may be an artefact or a strategic consequence of self-paced RSVP reading tasks. Because quite a few repetitive manual responses (i.e., button presses) are required to progress through the texts, readers are essentially forced to spend more time on the task at hand. As an advantageous epiphenomenon their understanding of the texts may improve. In addition to this inherent property of the task, readers may regulate their processing strategy for self-paced RSVP texts more actively. Because they cannot reread prior sentences or prior words, they are stimulated to consider the available visible input more carefully and update their mental representation of the text more frequently. Moreover, a sentence-by-sentence presentation mode may encourage readers to spend more time on sentence wrap up, to settle comprehension problems and integrate the sentence into the prior context before moving on to the next sentence.

This latter view on our results is particularly interesting considering a recent study by Tiffin-Richards and Schroeder (2018). They observed that German-speaking children in Grades 2 and 3 need to rely more on sentence-final wrap up than do older readers, presumably because less efficient readers postpone integrational processing to sentence boundaries. In Grade 4 the wrap-up processes of the German children exhibited a more adult-like pattern. These results present a rationale for why a sentence-by-sentence presentation mode reflects an optimal layout for our population of second- and third-grade pupils. Because these readers still heavily rely on sentence-final wrap up to obtain an integrated representation of a text, stimulating them to do so may increase the quality of their mental representation of a text considerably.

6.2. Who will benefit the most from segmented texts?

A common hypothesis is that less efficient readers will benefit the most from texts presented in segments (Castelhano & Muter, 2001; Chen, 1986; Lemarié et al., 2008). On the assumption that many pupils in Grades 2 and 3 are less efficient readers compared to older children and adults, our results are in line with this hypothesis. After all, the vast majority of our participants displayed a comprehension advantage for segmented texts, whereas prior studies did not reveal this pattern for older, more experienced readers

(Benedetto et al., 2015; Lemarié et al., 2008; Rayner et al., 2016; Van Silfhout, 2014; Van Silfhout, Evers-Vermeul, Mak, et al., 2014; Van Silfhout, Evers-Vermeul, & Sanders, 2014; but see Chung-Fat-Yim et al., 2017). Perhaps more interestingly, the hypothesis also predicts that less efficient readers within a population of beginning readers should benefit the most. Our exploratory analyses on the interplay of text layout and readers' decoding and comprehension skills showed that this holds to some extent. We did not observe a consistent pattern across experiments, however.

On the one hand, several findings confirmed the hypothesis. In Experiment 4, struggling second-grade readers displayed increased comprehension for texts in which saccadic eye movements were kept to a minimum. Proficient second-grade readers showed the opposite. Their understanding of a text decreased when saccadic eye movements were marginalized during reading. Furthermore, in Experiment 2 struggling second-grade readers processed expository texts that were presented sentence by sentence more quickly than expository texts that were presented in a traditional, continuous manner. This reading time advantage was not observed for more proficient readers, which again is consistent with the idea that struggling readers will benefit the most from a segmented layout.

On the other hand, it also became clear that in some specific situations segmented texts yield more advantages for proficient beginning readers. For example, in Experiment 4 proficient second-grade pupils obtained relatively high comprehension scores for texts when presented in a word-by-word manner – possibly because they were stimulated to adopt a more incremental processing strategy. Furthermore, the most striking evidence against the hypothesis that segmented texts solely assist more vulnerable beginning readers was obtained in Experiment 1. Proficient second-grade readers displayed a comprehension advantage (and a nonsignificant reading time advantage) for a discontinuous layout in which each sentence started on a new line of the page. The results for struggling second-grade readers showed no signs of these benefits. In fact, struggling readers slowed down while reading discontinuous texts. Hence, although the main purpose of a discontinuous layout is to assist more vulnerable beginning readers, the results of Experiment 1 enforce an unanticipated conclusion: not struggling beginning readers will benefit from a discontinuous layout, but proficient beginning readers will – reflecting a 'Matthew effect' (e.g., Shaywitz et al., 1995) for this reading intervention.

At this point we should reemphasize that our analyses on the interplay of text layout and pupils' reading skills were exploratory. Due to the nature of the data several separate analyses were conducted, thereby inflating the probability of reporting on false-positive findings. Furthermore, no consistent picture emerged when comparing the modulating effects of decoding skills to the modulating effects of comprehension skills. We therefore refrain from committing to an overly detailed interpretation of these exploratory analyses. The results allow for several interesting general conclusions, nevertheless. First, many pupils in Grades 2 and 3 will benefit from a segmented presentation mode of texts. Second, both reader

characteristics and text genre determine the efficacy of the presentation mode of a text. Third, severely segmented (i.e., word-by-word) texts have a beneficial influence on comprehension in some situations, even for more proficient beginning readers. Fourth, the potential advantages of severely segmented texts are most relevant in the early stages of reading acquisition. By the time pupils are attending Grade 3 the advantages seem less prominent, and a mildly segmented (i.e., sentence-by-sentence) mode becomes a more suitable way of presenting texts to most readers.

6.3. Digital approaches to enhance reading: The educational toolbox of future reading

The present study has several interesting implications for reading in educational environments. Perhaps the most remarkable practical implication is that segmented texts present a feasible alternative to texts with a traditional layout. Educational practitioners and developers should be made more aware of this as it will provide a tool for designing optimal learning environments, especially in the current digital era. To further illustrate the wide scope of possibilities offered by segmented texts and other powerful features of the ‘educational toolbox of future reading’, we will discuss the main findings of our study in light of several high-potential digital reading applications (i.e., SLTR, Spritz, BeeLine Reader, WebClipRead).

The computer-based reading intervention Span Limiting Tactile Reinforcement (SLTR) is arguably the most comprehensive method that will be addressed. SLTR was originally developed by Schneps and his colleagues to optimize the learning outcomes of dyslexic readers (Schneps, 2015; Schneps, Thomson, Sonnert, et al., 2013; Schneps et al., 2010; Schneps, Thomson, Chen, et al., 2013). In SLTR a digital text is reformatted into a single column with only a few words per line. Readers are encouraged to keep their gaze at the top of the page and control their reading pace with a jockey wheel to manually progress through a text one line at a time. Schneps et al. observed that while SLTR improves the reading of dyslexics, the benefits vary with individuals. They speculated, therefore, that the advantages of SLTR should carry over to more typically developing readers.

Several of the design principles of SLTR are rooted in the layout features that were assessed in the current study. First, short lines of text are presented. By adopting this presentation mode, the reader will profit from the word-preview effect, yet the overall distractibility of surrounding text is decreased. According to Schneps et al. this will diminish the incidence of redundant regressive eye movements. Second, by instructing readers to maintain their focus on the top line and scroll through the text line by line, SLTR closely resembles our self-paced sentence-by-sentence presentation mode. Consequently, SLTR may relax the demands on oculomotor control while promoting incremental integration, just like we postulated for a single-sentence presentation mode. Another beneficial feature of SLTR is that the finger action required to proceed to the next line of the text provides ‘tactile-kinesthetic’ reinforcement. This reinforcement may increase comprehension as it punctuates each phrase while reading (Schneps et al., 2010). Interestingly,

these potential benefits of manual reinforcement would provide an explanation for why we observed no overall comprehension advantage for the discontinuous layout in Experiment 1 (where each sentence was presented on a new line) yet observed a comprehension advantage for the single-sentence RSVP layout in Experiment 2. These two presentation formats share many design principles, yet only the latter provides tactile-kinesthetic support as readers press a button to proceed to the next sentence.

Whereas Schneps and his colleagues put their reading application to the empirical test, many developers fail to do that, or claim too much based on the available evidence. Spritz (www.spritzinc.com) is one of the applications that has been criticized for doing the latter and, perhaps as a result thereof, received a lot of attention in the media and academic world (Benedetto et al., 2015; Rayner et al., 2016; Ricciardi & Di Nocera, 2017; Schotter et al., 2014). In Spritz, single-word stationary RSVP reading is combined with an algorithm to highlight the letter that reflects the optimal viewing position of each word. Furthermore, longer words are displayed for a longer time and longer sentences have longer pauses at the end of the sentence. Although the criticism of several scholars concerning the bold claims of the developers is accurate (e.g., “only 20% of the reader’s time is spent processing content while 80% of the time is spent moving the eyes” and “every saccade has a penalty in both time and comprehension”, see Rayner et al., 2016), we should nuance the criticism somewhat. That is, based on the findings in the current study it is clear that Spritz incorporates several powerful design principles to assist beginning and other less efficient readers (e.g., segmented layout, additional time for sentence-final wrap-up). Hence, it would be premature to disregard this high-potential reading application.

The layout features of texts presented with methods like SLTR and Spritz are radically different from those of traditional texts. Instead, BeeLine Reader (www.beelinereader.com) is an interesting digital reading aid that modifies the appearance of texts in a subtle way. With BeeLine Reader the lines of a text are colored in a gradient by adopting a procedure that ensures that the end of one line and the beginning of the next are colored similarly. This is thought to facilitate visual tracking and it allows for smooth return sweeps. In fact, according to its developers “Beeline Reader is a research-backed tool that improves reading ability for students of all ages and skill levels”. Although we were unable to locate the research that would substantiate these claims, we agree with the developers and Schneps (2015) that BeeLine Reader is a clever application that has the potential to make reading more efficient, in particular for beginning readers who struggle to guide their eye gaze and attention from one line to the next.

Without a doubt, the current digital era produced several high-potential digital applications to streamline a wide range of reading processes – the reading applications discussed above only represent a small sample of the total number of potentially useful applications. In the long run, we envision even more comprehensive digital systems that take full advantage of the ongoing developments by combining the strengths of multiple reading applications. These systems of the future should be adaptive, meaning that

they will be able to monitor readers' eye-movement behavior, cognitive load, and text comprehension on the fly, and modify the layout of a text accordingly (cf., Öquist & Goldstein, 2003; Rosch & Vogel-Walcutt, 2013).

Before these adaptive systems can become operational, however, a more complete picture should be obtained of how specific adjustments to the layout of a text will affect reading. That is, although we believe that our study presents a significant step forwards in this endeavor, many open issues remain. For example, our data show that only proficient beginning readers (not struggling beginning readers) will benefit from a discontinuous layout with each sentence starting on a new line of the same page. At this point, we do not fully understand why that is the case. Furthermore, whereas we highlighted the beneficial effects of single-word and single-sentence presentation modes, Schneps et al. (2010) proposed that each line in a segmented text should contain two or three words instead. Moreover, Walker et al. (2005) developed a technique to make a segmented multiple-word layout even more powerful by using Visual Syntactic Text Formatting (VSTF). This procedure transforms block-shaped text into cascading patterns to help readers identify the grammatical structure of a sentence (VSTF is available as WebClipRead at www.liveink.com). Together, these design principles raise the question of what would constitute the optimal parsing procedure for a segmented text in terms of the length and content of its chunks (cf. Sharmin, Špakov, & Rähä, 2012; Young, 1984). Another issue concerns our interpretation of the increased reading times for segmented texts. We postulated that they reflect a more intensive, effortful processing strategy. However, other factors may play a role here. For example, most children will be unfamiliar with sentence-by-sentence or word-by-word texts and increased processing times for these presentation modes can thus be attributed to a novelty effect. In fact, one could argue that the pupils in our study were quite attracted to the layout of segmented texts – thereby increasing their attention – which in turn could affect their level of engagement and reading motivation. This would raise the question of whether segmented texts have a durable influence on children's reading approach, as novelty effects tend to wear off quickly (Reinders, 2017).

In all, the issues that were addressed in the present study illustrate that, in a way, we are still beginning to uncover the wealth of opportunities offered by information technology. Consequently, we should commit to investigate the potential benefits of alternate digital reading methods in a more systematic way. An important next step is to examine the collective influence of a set of prominent layout features (smaller vs. larger chunks of text, single- vs. multiple-line layout, self-paced vs. non-self-paced reading, manual reinforcement, gradient coloring, etc.) in series of well-controlled (longitudinal) studies in which the configuration of features is varied systematically among different age groups and types of readers. Only then we will be able to decipher how layout features and the characteristics of readers interact and make grounded decisions on effective text design.

6.4. Conclusion

Whether, when, and how segmented texts will be applied in classrooms or other educational settings will be closely related to the goals set by reading education. On the one hand, presenting texts with simplified layouts to beginning readers may prevent them from optimizing their reading skills for texts with a traditional layout. On the other hand, in the current day and age the landscape of reading is expanding at a remarkable rate. Many texts are read in a digital form, not seldom from small, handheld devices. Although this development induces substantial challenges to readers, it also presents many opportunities to improve the reading experience of people. Moreover, it suggests that presenting texts on paper in a continuous format is something of the past.

Regardless of whether traditionally formatted texts will disappear (gradually), we can conclude that segmented texts present a powerful tool to increase the reading abilities of beginning readers. Pending relevant data to suggest otherwise, a presentation mode in which texts appear one sentence at a time seems to combine the better of two worlds. First, while preserving word-preview benefits, disruptive return sweeps and long-distance regressions are removed from the eye-movement behavior. Second, a sentence-by-sentence presentation mode may promote sentence wrap-up and other integration processes, thereby improving the quality of reader's mental representation of a text. Together with the design principles of high-potential reading applications (SLTR, Spritz, BeeLine Reader, WebClipRead) these findings present a fruitful starting point to develop adaptive applications, aimed at augmenting the comprehension abilities of young, beginning readers.

7. References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious Mixed Models. <https://doi.org/arXiv:1506.04967>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using {lme4}. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- BeeLine Reader: making reading on-screen easier and faster. (n.d.). Retrieved May 18, 2018, from <http://www.beelinereader.com/>
- Benedetto, S., Carbone, A., Pedrotti, M., Le Fevre, K., Bey, L. A. Y., & Baccino, T. (2015). Rapid serial

- visual presentation in reading: The case of Spritz. *Computers in Human Behavior*, 45, 352–358.
<https://doi.org/10.1016/j.chb.2014.12.043>
- Best, R. M., Floyd, R. G., & Mcnamara, D. S. (2008). Differential Competencies Contributing to Children's Comprehension of Narrative and Expository Texts. *Reading Psychology*, 29(2), 137–164.
<https://doi.org/10.1080/02702710801963951>
- Busler, J. N., & Lazarte, A. A. (2017). Reading time allocation strategies and working memory using rapid serial visual presentation. *Journal of Experimental Psychology: Learning Memory and Cognition*, 43(9), 1375–1386. <https://doi.org/10.1037/xlm0000392>
- Cain, K., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing*, 11(5/6), 489–503. <https://doi.org/10.1023/A:1008084120205>
- Castelhano, M. S., & Muter, P. (2001). Optimizing the reading of electronic text using rapid serial visual presentation. *Behaviour & Information Technology*, 20(4), 237–247.
<https://doi.org/10.1080/01449290110069400>
- Chen, H.-C. (1986). Effects of reading span and textual coherence on rapid-sequential reading. *Memory & Cognition*, 14(3), 202–208. <https://doi.org/10.3758/BF03197693>
- Chung-Fat-Yim, A., Peterson, J. B., & Mar, R. A. (2017). Validating self-paced sentence-by-sentence reading: story comprehension, recall, and narrative transportation. *Reading and Writing*, 30(4), 857–869. <https://doi.org/10.1007/s11145-016-9704-2>
- Clifton, C., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. *Eye Movements: A Window on Mind and Brain*, 341–372.
- Drummond, A. (2013). Ibx farm. *Online Server: Http://Spellout. Net/Ibxfarm*.
- Evers-Vermeul, J., & Land, J. (2011). Short sentences easy to read? Effects of connectives on text comprehension by beginning readers. Paper presented at the 33rd Annual Conference of the Deutsche Gesellschaft für Sprachwissenschaft. Göttingen, 23 February 2011.
- Ferreira, V., Ferreira, F., & Henderson, J. M. (2015). Keith Rayner (1943–2015). *American Psychologist*, 70(6), 568. <https://doi.org/10.1037/a0039152>
- Fox, J. (n.d.). Effect Displays in R for Generalised Linear Models. Retrieved from <http://psfaculty.ucdavis.edu/bsjjones/effectdisplays.pdf>
- Fox, J., & Hong, J. (2009). Effect Displays in R for Multinomial and Proportional-Odds Logit Models: Extensions to the effects Package. *JSS Journal of Statistical Software*, 32(1). Retrieved from <http://www.jstatsoft.org/>
- Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression* (Second). Thousand Oaks {CA}: Sage. Retrieved from <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: from eye fixations to comprehension.

- Psychological Review*, 87(4), 329.
- Koornneef, A. W., & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, 54(4), 445–465. <https://doi.org/10.1016/j.jml.2005.12.003>
- Land, J. (2009). *Zwakke lezers, sterke teksten? : effecten van tekst- en lezerskenmerken op het tekstbegrip en de tekstwaardering van vmbo-leerlingen*. Eburon. Retrieved from <https://dspace.library.uu.nl/handle/1874/32240>
- Land, J., Sanders, T., & Van den Bergh, H. (2008). Effectieve tekststructuur voor het vmbo: een corpus-analytisch en experimenteel onderzoek naar tekstbegrip en tekstwaardering van vmbo-leerlingen voor studieteksten. *Pedagogische Studiën*, 82(2), 76–94. Retrieved from https://pure.uva.nl/ws/files/1085403/62360_301087.pdf
- Lemarié, J., Eyrolle, H., & Cellier, J.-M. (2008). The segmented presentation of visually structured texts: Effects on text comprehension. *Computers in Human Behavior*, 24(3), 888–902. <https://doi.org/10.1016/J.CHB.2007.02.016>
- Levasseur, V. M., Macaruso, P., Palumbo, L. C., & Shankweiler, D. (2006). Syntactically cued text facilitates oral reading fluency in developing readers. *Applied Psycholinguistics*, 27(3), 423–445. <https://doi.org/10.1017.S0142716406060346>
- Live Ink | - the ignition for cognition. (n.d.). Retrieved May 18, 2018, from <http://www.liveink.com/>
- McMaster, K. L., den Broek, P. van, A. Espin, C., White, M. J., Rapp, D. N., Kendeou, P., ... Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences*, 22(1), 100–111. <https://doi.org/10.1016/J.LINDIF.2011.11.017>
- Öquist, G., & Goldstein, M. (2003). Towards an improved readability on mobile devices: evaluating adaptive rapid serial visual presentation. *Interacting with Computers*, 15(4), 539–558. [https://doi.org/10.1016/S0953-5438\(03\)00039-0](https://doi.org/10.1016/S0953-5438(03)00039-0)
- Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, 11(10), 1129–1135. <https://doi.org/10.1038/nn.2187>
- R Core Team. (2017). R: A Language and Environment for Statistical Computing. Vienna, Austria. Retrieved from <https://www.r-project.org/>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372.
- Rayner, K., & Clifton, C. (2009). Language processing in reading and speech perception is fast and incremental: implications for event-related potential research. *Biological Psychology*, 80(1), 4–9. <https://doi.org/10.1016/j.biopsycho.2008.05.002>

- Rayner, K., Schotter, E. R., Masson, M. E. J., Potter, M. C., & Treiman, R. (2016). So Much to Read, So Little Time. *Psychological Science in the Public Interest*, 17(1), 4–34.
<https://doi.org/10.1177/1529100615623267>
- Reinders, H. (2017). Digital Games and Second Language Learning. In S. Thorne & S. May (Eds.), *Language, Education and Technology* (pp. 1–15). Cham: Springer International Publishing.
https://doi.org/10.1007/978-3-319-02328-1_26-2
- Ricciardi, O., & Di Nocera, F. (2017). Not so fast: A reply to Benedetto et al. (2015). *Computers in Human Behavior*, 69, 381–385. <https://doi.org/10.1016/j.chb.2016.12.047>
- Rosch, J. L., & Vogel-Walcutt, J. J. (2013). A review of eye-tracking applications as tools for training. *Cognition, Technology & Work*, 15(3), 313–327. <https://doi.org/10.1007/s10111-012-0234-7>
- Schneps, M. (2015). Using Technology to Break the Speed Barrier of Reading. *Scientific American MIND*. Retrieved from <https://www.scientificamerican.com/article/using-technology-to-break-the-speed-barrier-of-reading/>
- Schneps, M. H., O’Keeffe, J. K., Heffner-Wong, A., & Sonnert, G. (2010). Using Technology to Support STEM Reading. *Journal of Special Education Technology*, 25(3), 21–33.
<https://doi.org/10.1177/016264341002500304>
- Schneps, M. H., Thomson, J. M., Chen, C., Sonnert, G., & Pomplun, M. (2013). E-Readers Are More Effective than Paper for Some with Dyslexia. *PLoS ONE*, 8(9), e75634.
<https://doi.org/10.1371/journal.pone.0075634>
- Schneps, M. H., Thomson, J. M., Sonnert, G., Pomplun, M., Chen, C., & Heffner-Wong, A. (2013). Shorter Lines Facilitate Reading in Those Who Struggle. *PLoS ONE*, 8(8), e71161.
<https://doi.org/10.1371/journal.pone.0071161>
- Schotter, E. R., Tran, R., & Rayner, K. (2014). Don’t Believe What You Read (Only Once). *Psychological Science*, 25(6), 1218–1226. <https://doi.org/10.1177/0956797614531148>
- Sharmin, S., Špakov, O., & Rähä, K.-J. (2012). The Effect of Different Text Presentation Formats on Eye Movement Metrics in Reading. *Journal of Eye Movement Research*, 5(3), 1–9.
<https://doi.org/10.16910/jemr.5.3.3>
- Shaywitz, B. A., Holford, T. R., Holahan, J. M., Fletcher, J. M., Stuebing, K. K., Francis, D. J., & Shaywitz, S. E. (1995). A Matthew Effect for IQ but Not for Reading: Results from a Longitudinal Study. *Reading Research Quarterly*, 30(4), 894. <https://doi.org/10.2307/748203>
- Spritz. (n.d.). Retrieved May 18, 2018, from <http://spritzinc.com/>
- Tiffin-Richards, S. P., & Schroeder, S. (2018). The development of wrap-up processes in text reading: A study of children’s eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, Advance on*, 1–13. Retrieved from <http://psycnet.apa.org/record/2018-09033-001>

- Van Renswoude, D. R., Johnson, S. P., Raijmakers, M. E. J., & Visser, I. (2016). Do infants have the horizontal bias? *Infant Behavior and Development, 44*, 38–48.
<https://doi.org/10.1016/j.infbeh.2016.05.005>
- Van Silfhout, G. (2014). *Fun to read or easy to understand? Establishing effective text features for educational texts on the basis of processing and comprehension research*. Utrecht: LOT. Retrieved from http://www.lotpublications.nl/Documents/368_fulltext.pdf
- Van Silfhout, G., Evers-Vermeul, J., Mak, W. M., & Sanders, T. J. M. (2014). Connectives and layout as processing signals: How textual features affect students' processing and text representation. *Journal of Educational Psychology, 106*(4), 1036–1048. <https://doi.org/10.1037/a0036293>
- Van Silfhout, G., Evers-Vermeul, J., & Sanders, T. J. M. (2014). Establishing coherence in schoolbook texts: How connectives and layout affect students' text comprehension. *Dutch Journal of Applied Linguistics, 3*(1), 1–29. <https://doi.org/10.1075/dujal.3.1.01sil>
- Walker, S., Schloss, P., Fletcher, C. R., Vogel, C. A., & Walker, R. C. (2005). Visual-syntactic text formatting: A new method to enhance online reading. *Reading Online*. Retrieved from http://www.liveink.com/VSTF_ReadingOnline_IRA_2005_Walker.pdf
- Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160-168.
<https://doi.org/10.1016/j.tics.2011.02.005>
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis* Springer-Verlag. *New York*. Retrieved from https://scholar.google.nl/scholar?q=++H.+Wickham.+ggplot2%3A+Elegant+Graphics+for+Data+Analysis.+Springer-Verlag+New+York%2C+2009.&btnG=&hl=nl&as_sdt=0%2C5
- Young, S. R. (1984). RSVP: A task, reading aid, and research tool. *Behavior Research Methods, Instruments, & Computers, 16*(2), 121–124. <https://doi.org/10.3758/BF03202369>
- Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., ... Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences of the United States of America, 109*(28), 11455–9.
<https://doi.org/10.1073/pnas.1205566109>