# Syntactically cued text facilitates oral reading fluency in developing readers

VALERIE MARCIARILLE LEVASSEUR
*Haskins Laboratories and University of Connecticut*

PAUL MACARUSO
*Haskins Laboratories and Community College of Rhode Island*

LAURA CONWAY PALUMBO
*Haskins Laboratories*

DONALD SHANKWEILER
*Haskins Laboratories and University of Connecticut*

ADDRESS FOR CORRESPONDENCE
Valerie Marciarille LeVasseur, Department of Psychology, Community College of Rhode Island,
400 East Avenue, Warwick, RI 02886-1807. E-mail: vlevasseur@ccri.edu

ABSTRACT
Can fluency in oral reading be facilitated by formatting text to preserve major syntactic boundaries?
Seven-, 8-, and 9-year-old children read aloud passages under two text format conditions. In the
structure-preserving condition, the ends of lines coincided with ends of clauses; in the phrase-disrupting
condition, line breaks always interrupted a phrasal unit. Experiment 1 showed that oral reading fluency,
as indexed by skill in phrasal reading, was rated higher when children were reading in the structure-
preserving condition. In addition, the structure-preserving condition resulted in significantly fewer false
starts at the beginning of lines following a return sweep. The results of Experiment 2, in which texts of
varying levels of difficulty were read by slightly older readers, confirmed both findings. Measures of
fluency were correlated with other language and reading measures; however, no effects of format were
obtained on oral reading rate (words correct per minute). Taken as a whole, these findings indicate a
benefit of keeping clausal units intact in promoting fluent reading by facilitating the transition from
one line to the next.

Many researchers and teachers share a common belief that fluency in reading
connected material is a good indicator of overall competence in reading. Ideas
as to what should count as reading fluency vary, however. Some conceptions
place the emphasis solely on speed of word recognition (Meyer & Felton, 1999;
Torgesen, Rashotte, & Alexander, 2001). That emphasis is understandable in
view of the many indications that skill in word recognition and its phonological
underpinnings is so often the major limiting factors in reading comprehension

(Fletcher et al., 1994; Foorman, Francis, Novy, & Liberman, 1991; Shankweiler et al., 1999; Stanovich, Cunningham, & Feeman, 1984; Stanovich & Siegel, 1994).

However, there is more to fluency than word-level processing. One indication of this is the evidence that speed of reading words in text (words correct per minute [WCPM]) is more highly correlated with measures of comprehension than speed of reading individual words, unconnected in a list (Fuchs, Fuchs, Hosp, & Jenkins, 2001; Fuchs, Fuchs, & Maxwell, 1988). Text reading speeds reflect constraints posed by higher level language structure as well as characteristics of individual words. To read text with comprehension one needs to process both the individual words and to parse their phrasal groupings. In addition, it is apparent from listening to people read that there exists an ability above and beyond word recognition that contributes to naturalness in reading aloud. This is the ability to supply the appropriate prosody. Prosody includes suprasegmental aspects of speech such as sentence pitch contours, stress rhythms, and pauses at major syntactic breaks. Thus, one may consider dysfluent reading to include not only hesitations, stumbles, and occasional errors in identifying words, but also problems in phrasing, emphasis, and intonation.

Acknowledging the importance of these additional dimensions of fluency, the National Reading Panel Report (2000) defines reading fluency as the ability to "read text with speed, accuracy, and proper expression" (pp. 3–5). Reading teachers and diagnosticians are often aware of these dimensions and use instruments such as the Gray Oral Reading Tests (Weiderholt & Bryant, 2001) to assess phrasing and prosody as well as facility with word recognition. Although a speed/accuracy measure, usually reported as WCPM, embraces an important part of fluency, such a measure does not directly capture the success with which the reader organizes the succession of words into phrases with appropriate pauses and stress. These considerations have led us to examine both word and supraword aspects of fluency.

Given the long-standing belief that prosody is an important aspect of reading skill, there has been surprisingly little study of how prosodic skills develop. Models of reading commonly assume that proficient readers automatically segment the sentences of text into syntactically and semantically appropriate units or phrases (Gough, 1985; Just & Carpenter, 1987; Laberge & Samuels, 1985; Rumelhart, 1985). On the simplest assumption about the relation between reading and speech, sentence parsing skills gained from experience with spoken language would transfer automatically to reading once word recognition skills are in place. Observations of children reading aloud suggest that this assumption is incorrect. Parsing into phrasal and clausal units seems to be a problem for many developing readers (Chomsky, 1976; Fuchs et al., 2001; Rasinski, 1994; Schreiber, 1980, 1991). Therefore, we should entertain the possibility that parsing of printed material may require a separate learning step over and above the rapid identification of individual words (Schreiber, 1980).

One reason that learning to parse text may present a challenge is because, unlike speech, printed text conveys few direct prosodic cues. Punctuation marks offer only a rough and somewhat unreliable guide to prosody. In speech, prosody serves as a cue to the grouping of words into syntactic and semantic units and aids the listener in apprehending the message. In fact, children seem to be especially reliant on prosodic cues in parsing speech (Schreiber, 1991; Schreiber & Read, 1980). When

confronted with text, however, where prosodic cues are mostly lacking, children may have difficulty imposing the necessary phrasal structure. Thus, some children may have adequate word recognition and listening comprehension abilities, and nonetheless struggle to become fluent readers because they have underdeveloped text parsing skills (Rasinski, 1994).

Schreiber (1980, 1991) has suggested that the acquisition of text parsing skills may encounter an impediment relating to the arrangement of text on the page. As children move to texts with complex sentences, requiring more than one line, they must learn to maintain rate and expression across breaks at the ends of lines. These breaks often disrupt structural units. The following is an excerpt from a text intended for second graders (Almada, 1997):

> I know that you would have rather stayed
> at home this vacation, but being alone is just
> too dangerous. Its safer for you to be at
> Grandpa's house. You will have fun.

Note how phrasal structure is interrupted by line breaks. Consider, for example, the prepositional phrase "at Grandpa's house." This phrase should be parsed as a unit; however, because it is interrupted by a line break, parsing can only be completed after a sweep of the eyes to the next line. Especially for an inexperienced reader, the interruption in processing necessitated by a return sweep could be disrupting. Normally progressing readers make the transition from word by word reading to phrasal reading at some point after the first grade (Chall, 1983; Dowhower, 1987). At roughly the same time that young readers are beginning to group words into phrases, they must also learn to accommodate frequent line breaks. Arguably, it is difficult for them to coordinate the several processes required to maintain fluent reading across line breaks. Providing support at this stage by adjusting the location of line breaks to align with sentence structure may prove beneficial to fluent, phrasal reading.

In fact, the literature reports several attempts by researchers to study the effects on reading fluency of altering the conventional layout of text on a page, some by inserting extra spaces to separate syntactic phrases (Bever, Jandreau, Burwell, Kaplan, & Zaenen, 1990; Carver, 1970; Jandreau & Bever, 1992; O'Shea & Sindelar, 1983) and others by utilizing line breaks to mark phrase boundaries (Graf & Torrey, 1966; Marciarille-LeVasseur, Macaruso, & Shankweiler, 2001; Wood, 1975). Some of these studies have observed a speed advantage when children or adults read text that was segmented to keep syntactic phrases intact (see Bever et al., 1990; Jandreau & Bever, 1992; Mason & Kendall, 1979; Wood, 1975).

In the present study, we consider the effects on reading fluency (speed, phrasing and prosody) of altering line breaks so that they preserve syntactic structures. We also consider whether these manipulations have an impact on reading comprehension. Although there is evidence of an association between reading fluency, as indexed by WCPM, and reading comprehension (Fuchs et al., 2001), the presumed link between prosodic reading and reading comprehension is more tentative. One finding favoring such a link comes from a study by Young and Bowers (1995), who found a positive relationship between higher fluency ratings, which captured

syntactic–semantic phrasing, and reading comprehension in older children. However, experiments evaluating the effects of text segmentation on reading comprehension have led to mixed results, both in studies of children and adults. Although some text manipulations have led to benefits in reading comprehension (Bever et al., 1990; Jandreau & Bever, 1992; O'Shea & Sindelar, 1983; Weiss, 1983), in other cases similar manipulations have been ineffective (Carver, 1970; Marciarille-LeVasseur et al, 2001; Wood, 1975).

We designed the present study primarily to test the effects of syntactically segmented text on *oral* reading fluency in developing readers. Studies that examine reading aloud make it possible to identify specific aspects of fluency that are masked in silent reading, namely the phrasing and prosodic elements. We contrasted oral reading of prepared text in which clausal structure is kept intact at line breaks with oral reading of text in which line breaks disrupt clausal and phrasal structure. We predicted that visual cuing of syntactic structure would aid students in grouping syntactic units in text and thus promote fluency. Such benefits might also be expected to facilitate comprehension.

## EXPERIMENT 1

### Method

*Participants.* The sample consisted of 35 children (19 males, 16 females) drawn from two urban elementary schools in Rhode Island. Some of the children were second graders (7- and 8-year-olds) tested at the end of the school year, and others were third graders (8- and 9-year-olds) tested at the beginning of the next year. All were native speakers of English. Three children were removed because they read too slowly or too inaccurately, such that their average WCPM on the experimental passages fell below 50 (the criterion also adopted by Dowhower, 1987). Analyses were based on the remaining 32 participants (23 second graders, 9 third graders; 16 males, 16 females) with an average age of 8 years 2 months. Participants received colored pencils for participating.

*Materials.* Passages were adapted from a children's text, *Oscar and Tatiana* (Almada, 1997), which was chosen for this study because it has a lexile value of 300, deemed appropriate for second graders. Lexile scores are based on a readability formula that matches readers with appropriate texts, taking into account word difficulty and sentence complexity (Schnick & Knickelbine, 2000). The text comprised a series of letters between two children and their mother. The first four letters, averaging 102 words each, were selected as the test passages. The materials were reproduced in two formats equated for mean words per line (6) and mean lines per passage (18): (a) a structure-preserving (SP) condition in which each end of a line corresponded to a clause boundary, and (b) a phrase-disrupting (PD) condition in which the end of a line interrupted a phrasal constituent. The test passages were preceded by a practice passage that introduced the characters in the text. The practice passage was formatted like the published text such that clausal structure was preserved at some line breaks and disrupted at others (see Appendix A).

The four test passages were shown in the same order (Passages 1, 2, 3, 4) to keep the story content intact. The SP and PD formatted versions of these passages were counterbalanced. Four orders of the versions were created. For example, participants in Order 1 received Passage 1 in SP form, Passage 2 in PD form, Passage 3 in SP form, and Passage 4 in PD form. The number of participants in each order was: seven in Order 1, eight in Order 2, nine in Order 3, and eight in Order 4.

*Procedure.* Data were collected from students' oral reading of the experimental passages and from a series of language and reading skill tasks. Participants took part in two sessions: a group session, and an individual session. During the 10-min group session, children were asked to write to dictation 50 words from the Developmental Spelling Analysis (DSA; Ganske, 1999). The DSA is an inventory that assesses students' knowledge of important orthographic features. We used lists covering spelling patterns typically taught in the second grade.

The individual session lasted approximately 15 min. During this session, the participant was seated at a table beside the tester and in front of a microphone. He or she was first instructed to read the practice passage aloud as well as possible. This was followed by two comprehension questions that were read aloud by the tester and responded to orally by the participant. A participant who responded incorrectly to both questions during practice was instructed to read more carefully. This occurred only twice. Next, the participant viewed the first experimental passage and was instructed to read it aloud. Four comprehension questions followed (see Appendix A). The same procedure was used for the remaining three passages. No feedback was provided on experimental questions. This portion of the session was recorded on audiotape.

At the end of the individual session, a measure of expressive vocabulary, the Critchlow Vocabulary Test (Critchlow, 1996), was administered. This list consists of 75 words of increasing difficulty (e.g., hot, imprisoned). Upon hearing each word, the participant responded aloud with its antonym (e.g., cold, free). This procedure continued until the participant offered five consecutive wrong responses. Students in the second grade (ages 7–8 years) typically score 13–17 correct.

In addition, the Gates–MacGinitie Reading Test, Fourth Edition, Level 2 (MacGinitie et al., 2000), had been administered by the school system at the end of the second grade. Grade equivalencies and percentile ranks from the comprehension subtest of the Gates–MacGinitie Reading Test were used in this study.

*Scoring.* Six experimental measures were obtained from the audiotapes. Two of the measures relate to speed and accuracy.

1. *WCPM* and
2. percentage of *word errors*, which include decoding errors and word omissions. Percentages were obtained by dividing the number of word errors by the number of words in the text.

Three measures were related to prosody.

3. A global *fluency rating* for each passage. The ratings were based on a scale adapted from the National Association of Educational Progress (National Center for Education Statistics, 1995). The fluency rating scale gives the greatest weight to proper phrasing in reading aloud. A 4-point scale was used in which 1 indicates that phrasing was mainly absent (reading in a listwise, word by word manner) and 4 indicates appropriate phrasing (reading in syntactically organized phrasal groups). The 2 and 3 ratings were assigned to intermediate cases. Two raters listened to the audiotapes and provided a rating for each passage. Both raters are teachers experienced with children's reading. Upon listening to the children read the text aloud, it was noticed that some children were hesitating and stumbling at line breaks. These stumbles seemed to be contributing to disruptions in prosody so we decided to analyze them separately, calling them "false starts." We also analyzed hesitations and stumbles that occurred elsewhere in the text. These are called "other dysfluencies."

4. The percentage of *false starts,* which include hesitations or stumbles on the first word of a line, rereading the first word of a line after a stumble within the line, and rereading the end of the previous line. For example, consider the following text:

> We are all fine today. Yesterday Tammy got in trouble when she tripped over a paint can . . .

If the participant read this as, "We are all fine today. Yesterday Tammy got in trouble. . . . Tammy got in trouble when she tripped over a paint can," this was scored a false start. Reading it as, "We are all fine today. Yesterday Tammy got a . . . in . . . in trouble when she tripped over a paint can . . ." was also a false start. Percentages were obtained by dividing the number of lines containing a false start by the number of lines in the text.

5. The percentage of *other dysfluencies*, which include hesitations within a line, and stumbles on or rereading any word other than the first word of a line. Percentages were obtained by dividing the number of lines containing an *other dysfluency* by the number of lines in the text.

The final measure assesses comprehension.

6. The percentage of correct responses on the *comprehension* questions.

*Reliability.*   A reliability index, based on a subset of the readings, was obtained for false starts, other dysfluencies, and word errors. This subset consisted of one SP and one PD passage from 21 participants, randomly chosen. The first and third authors scored the audiotapes independently. Agreement was 92% for false starts, 88% for other dysfluencies, and 86% for word errors. Scores based on the first author's judgments were used in analyses.

To assess internal consistency, Cronbach's alpha was computed for each measure based on performance across the four passages (two SP, two PD). Alpha values for WCPM, word errors, and fluency ratings were above .85, indicating

Table 1. *Effect of text format on measures of fluency and comprehension for Experiment 1*

| | Mean | | Standard Deviation | |
| --- | --- | --- | --- | --- |
| Measures | Structure Preserving | Phrase Disrupting | Structure Preserving | Phrase Disrupting |
| WCPM | 108 | 102 | 25.6 | 25.4 |
| Word errors (%) | 1.6 | 2.0 | 1.8 | 2.6 |
| Fluency ratings (scale 1–4) | 2.7 | 2.3 | 0.7 | 0.6 |
| False starts (%) | 1.7 | 5.8 | 2.2 | 4.4 |
| Other dysfluencies (%) | 10.3 | 10.3 | 7.2 | 6.7 |
| Comprehension correct (%) | 82 | 77 | 13.9 | 19.0 |

high internal consistency. The alpha values for false starts, other dysfluencies, and comprehension were in the low to moderate range at .56, .62, and .49, respectively.

### Results

We first present descriptive statistics for the language and reading skill measures (Gates–MacGinitie comprehension, vocabulary, spelling). Then, we report analyses examining the effects of text format (SP, PD) on the experimental measures obtained from reading the passages aloud (WCPM, word errors, fluency ratings, false starts, other dysfluencies, comprehension). Effect sizes were computed using Cohen's *d* (Cohen, 1988). Finally, bivariate correlations are examined for the experimental and skill measures.

*Language and reading skill measures.* The average percentile rank on the Gates–MacGinitie comprehension subtest was 69 ($SD = 17$), equivalent to a third-grade level. Participants averaged 28 words (out of 75) correct ($SD = 8$) on the Critchlow vocabulary test, indicating performance at a fifth-grade level. The average spelling score on the DSA was 38 (out of 50 words) correct ($SD = 11$).[1]

*Text format effects.* One-way analyses of variance (ANOVAs) revealed no order effects for any of the dependent variables, $Fs (3, 28) < 1.85$, $ps > .16$ for WCPM, false starts, other dysfluencies, word errors, fluency ratings, and comprehension. Hence, correlated *t* tests were computed to test for effects of format (SP vs. PD) on each dependent variable. Means are shown in Table 1. A significance level of .01 was set to control for inflated Type I error rate associated with multiple tests. Although participants read, on average, slightly more WCPM in the SP condition (108) than the PD condition (102), the effect of format on WCPM failed to reach significance, $t (31) = 2.52$, $p = .02$, $d = .22$. In addition, there was no significant difference between conditions in percentage of word errors, $t (31) = 1.52$, $p = .14$, $d = .20$.

Turning to the prosody measures, the two raters showed fair agreement for fluency ratings ($r = .71$, $p < .01$), and their ratings were averaged in the analyses.

Table 2. *Correlations among fluency, comprehension, and skill measures in Experiment 1*

| Measures | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 1. WCPM | — | −.61** | .59** | −.39* | −.58** | .13 | .38* | .71** | .45* |
| 2. Word errors | | — | −.31 | .47** | .45* | −.33 | −.46** | −.76** | −.22 |
| 3. Fluency ratings | | | — | −16 | −.35 | .30 | .06 | .37* | .12 |
| 4. False starts | | | | — | .63** | −.05 | −.34 | −.39* | −.10 |
| 5. Other dysfluencies | | | | | — | .21 | −.17 | −.45* | −.12 |
| 6. Comprehension | | | | | | — | .47** | .19 | .37* |
| 7. Vocabulary | | | | | | | — | .33 | .54** |
| 8. Spelling | | | | | | | | — | .23 |
| 9. Gates–MacGinitie comp. | | | | | | | | | — |

*$p < .05$. **$p < .01$.

There was a significant effect of format on fluency ratings, $t(31) = 3.19$, $p < .01$. Higher ratings were obtained for the SP condition than the PD condition. The effect size was moderate ($d = .48$). A significant effect of format was also obtained for percentage of false starts, $t(31) = 5.03$, $p < .01$. The effect size for false starts was strong ($d = 1.1$). Looking at raw numbers, false starts were reduced from a total of 69 in the PD condition to 31 in the SP condition. Percentage of *other dysfluencies* did not differ by format, $t(31) = 0.04$, $p = .97$, $d = .01$.

There was no significant effect of format on comprehension, $t(31) = 1.22$, $p = .23$. Although results favored the SP condition, the effect size was weak ($d = .28$).

*Correlations.* Table 2 shows correlations among the experimental measures (WCPM, word errors, fluency ratings, false starts, other dysfluencies, comprehension) and skill measures (vocabulary, spelling, Gates–MacGinitie comprehension). The experimental measures were averaged across format.[2]

Among the experimental measures, higher WCPM was associated with fewer word errors, false starts, and other dysfluencies, and with higher fluency ratings. Other significant correlations include word errors with false starts and other dysfluencies. Those who made more word errors also made more false starts and/or other dysfluencies. In addition, false starts correlated significantly with other dysfluencies. In sum, there is a moderate level of internal consistency among the fluency measures. However, none of the fluency measures correlated significantly with the experimental comprehension measure.

Predictably, the language and reading skill measures correlated significantly with WCPM. Participants with higher vocabulary, spelling, and/or Gates–MacGinitie comprehension scores showed higher WCPM. Those with higher vocabulary and spelling scores made significantly fewer word errors. Spelling also correlated significantly with false starts, other dysfluencies, and fluency ratings. Finally, vocabulary and Gates–MacGinitie comprehension correlated significantly with the experimental comprehension measure. As would be expected, students with higher vocabulary and/or Gates–MacGinitie comprehension scores showed better comprehension of the text materials.

*Discussion*

The chief findings of this study provide support for our guiding hypothesis: formatting text with line breaks that preserve syntactic structure enabled more fluent oral reading. Participants produced more phrasally organized readings, as indexed by fluency ratings, in the SP condition compared to the PD condition. It appears that the children made use of textual supports to more readily parse the text. The false starts measure supports and extends this claim. There was a marked reduction in the number of false starts in the SP condition compared with the PD condition. This result indicates that formatting text by adjusting line length to keep clausal structure intact facilitates the smooth transition from line to line. Because children appear to rely heavily on prosodic cues to parse speech (Schreiber, 1991), it is not surprising that adjusting line length to avoid phrasal interruptions would allow them to read with greater fluidity and expression. In terms of WCPM, there was a small benefit afforded by the SP condition, but this was not significant. There was no effect of text format on the comprehension measure. The latter finding is consistent with a number of others (e.g., Carver, 1970; Marciarille-LeVasseur et al., 2001; National Reading Panel, 2000; Wood, 1975).

The impact of syntactic formatting on phrasing and prosody was possible to assess because we chose to investigate oral reading rather than silent reading. Previous studies on text formatting have examined silent reading, and thus the results were limited to effects (or lack thereof) on WCPM and comprehension. Measures tied to oral reading, fluency ratings and false starts, turned out to be quite sensitive to text formatting; however, the latter measure showed less internal consistency than the former. Thus, in our second experiment, we decided to examine these measures further as they relate to oral reading fluency.

Accordingly, Experiment 2 was designed to (a) replicate and extend the effect of text formatting on measures of phrasing and prosody, including fluency ratings and false starts, and (b) examine the effects of text formatting on speed and accuracy measures (WCPM and word errors) with more challenging text and somewhat older children.

EXPERIMENT 2

In Experiment 2, all participants were third graders (8- and 9-year-olds), tested at the end of the school year. They read text materials containing more advanced vocabulary and longer and more complex sentences than those used in Experiment 1. In addition, we opted to assess comprehension with a story recall task rather than comprehension questions as in Experiment 1. Story recall is widely used in classroom settings.

*Method*

*Participants.* The sample consisted of 26 children attending two third-grade classes from a suburban public school in southeastern Massachusetts. There were 15 males and 11 females, all native speakers of English. Two children were excluded because of speech difficulties (stuttering), which significantly compromised

their oral reading skills, and one child was excluded because he produced greater than 5% word errors in reading the practice passage. Data from the remaining 23 participants (12 males, 11 females) with a mean age of 9 years and 0 months are reported. Participants received colored pencils for participating.

*Materials.* One fictional passage was taken from each of four children's texts. Two passages were excerpts from texts intended for third graders: *What a Day*! (Odgers, 1987) with a lexile value of 520, and *The Midnight Pig* (Krueger, 1997) with a lexile value of 500. The remaining two passages were taken from texts appropriate for fourth graders: *The Week of the Jellyhopper* (Cartwright, 1995) with a lexile value of 700, and *Glumly, USA* (Klein, 1995) with a lexile value of 700. The four passages averaged 268 words each. The passages were constructed in the same manner as Experiment 1; each was reproduced with an SP and PD form (see Appendix B).[3]

Two orderings were created such that each participant saw two passages in SP form and two in PD form. The two third-grade passages were presented first, followed by the two fourth-grade passages. One passage at each grade level was in SP form and the other in PD form. Twelve participants received Order 1 (Passage 1, SP; Passage 2, PD; Passage 3, PD; Passage 4, SP) and 11 participants received Order 2 (Passage 1, PD; Passage 2, SP; Passage 3, SP; Passage 4, PD).

*Procedure.* As in Experiment 1, data were collected from oral readings of the experimental passages and from a set of language and reading skill tasks. Participants again took part in two sessions: group session, and an individual session. During the 45-min group session, participants first received the DSA spelling test (see Experiment 1). They then completed the comprehension subtest of the Gates–MacGinitie Reading Test, Fourth Edition, Level 3 (MacGinitie et al., 2000).

The individual session lasted approximately 30 min. During this session, the participant was seated at a table beside the tester and in front of a microphone. A practice passage was administered first. This passage was chosen from a 500-lexile-level text (Dugan, 1988), unmodified from its conventional format. The participant was instructed to read the passage aloud as well as possible.

Once the practice passage was completed, the participant was given the first experimental passage and instructed to read it aloud as well as possible. Afterward, comprehension was assessed via free recall of story line and details. The participant was asked to "tell [the tester] what happened in the story." If recall was clearly incomplete, the tester asked, "Is there anything else?" No more than one prompt was offered for each test passage. The same procedure was used for all four passages. The oral reading and recall responses were audiotaped for later analysis. The Critchlow Vocabulary Test and the word attack subtest (consisting of 60 English-like nonwords) from the Woodcock–Johnson Psychoeducational Battery (Woodcock & Johnson, 1993) were administered at the end of the session.

*Scoring.* As in Experiment 1, six experimental measures were obtained from the audiotapes: two measures of reading speed and accuracy (WCPM, word errors), three measures of prosody (fluency ratings, false starts, other dysfluencies), and a measure of comprehension. For comprehension, the use of story recall led to a new dependent measure, *recall ratings*. This rating was based on two aspects

to story recall. First, the rater determined how well the participants understood the point of the story, the "gist." Second, the rater scored the number of specific details from the passage the participant could recall. From this information, the rater assigned a rating on a scale from 1 to 4 as follows: 1 = *misses the gist of the story*, 2 = *gets the gist but supplies limited supporting details* (0–4 details), 3 = *gets the gist and supplies some details* (5–10 details), and 4 = gets the gist and supplies many details (11–15 details).

*Reliability.*    A reliability index was obtained for false starts, other dysfluencies, and word errors as in Experiment 1. One SP and one PD passage from each of 10 randomly chosen participants were examined. Agreement among the two raters was 96% for false starts, 89% for other dysfluencies, and 83% for word errors. Again, scores based on the first author's judgments were used in analyses. To assess reliability for the recall ratings, the second author rated two-thirds of the recall samples. Agreement with the first author's ratings was 85%. The first author's ratings were used in analyses.

As in Experiment 1, Cronbach alphas were obtained for each experimental measure to determine internal consistency. Alpha values for WCPM, word errors, and fluency ratings were >.85, indicating high internal consistency. The alpha values for false starts, other dysfluencies, and comprehension were in the low to moderate range at .43, .68, and .56, respectively.

### Results

We first present descriptive statistics for the language and reading skill measures (Gates–MacGinitie comprehension, oral vocabulary, spelling, word attack). Then, we report analyses examining the effects of text format (SP, PD) on the experimental measures. Finally, bivariate correlations were examined for the experimental and skill measures.

*Language and reading skill measures.*    The average percentile rank on the Gates–MacGinitie comprehension subtest was 69 ($SD = 28$), equivalent to a fourth-grade level. Participants averaged 38 (out of 75) words correct ($SD = 8$) on the Critchlow vocabulary test, placing them at the seventh-grade level. Their average spelling score on the DSA was 46 (out of 50) words correct ($SD = 4$). Average performance on word attack was 33 (out of 60) items correct ($SD = 6$). This falls at a fifth-grade level. In sum, this group of children averaged above third-grade level on most measures, based on national US norms.

*Text format effects.*    One-way ANOVAs revealed no order effects for any variable, $F$s $(1, 21) < 2.08$, $p > .16$ for WCPM, word errors, fluency ratings, false starts, other dysfluencies, and recall ratings. Hence, correlated $t$ tests were computed to test for effects of format (SP vs. PD) on each dependent variable. A significance level of .01 was used for all tests. There was no effect of format on WCPM, $t (22) = .04$, $p = .97$, $d = .00$. As shown in Table 3, mean WCPM were nearly identical for the SP and PD conditions. As in Experiment 1, there was also no significant effect of format on word errors, $t (22) = 1.82$, $p = .08$, $d = .23$, although there was a trend favoring the SP condition.

Table 3. *Effect of text format on measures of fluency and comprehension for Experiment 2*

| | Mean | | Standard Deviation | |
| --- | --- | --- | --- | --- |
| Measures | Structure Preserving | Phrase Disrupting | Structure Preserving | Phrase Disrupting |
| WCPM | 123.3 | 123.4 | 30.8 | 33.7 |
| Word errors (%) | 2.0 | 3.1 | 1.9 | 2.6 |
| Fluency ratings (scale 1–4) | 2.8 | 2.5 | 0.7 | 0.7 |
| False starts (%) | 1.7 | 4.0 | 2.7 | 3.1 |
| Other dysfluencies (%) | 14.0 | 15.3 | 6.5 | 7.9 |
| Recall ratings (scale 1–4) | 2.3 | 2.2 | 1.0 | 0.8 |

The two raters showed strong agreement in their fluency ratings ($r = .85$, $p < .01$), so their ratings were averaged in the analyses. As in Experiment 1, there was a significant effect of format on fluency ratings, $t(22) = 2.73$, $p < .01$, $d = .41$. Participants read more phrasally in the SP condition than the PD condition.

When analyzing false starts, we encountered a difficulty that we could not anticipate from Experiment 1. In Experiment 1, we defined false starts as hesitations or stumbles on the first word of a line, rereading that word, or rereading the end of the previous line. Participants in Experiment 2 seemed to be following an error correction strategy that resulted in a scoring ambiguity. When they stumbled anywhere within a sentence, they would frequently return to the beginning of the sentence and reread (called a sentence onset return).[4] Given that many of the beginnings of lines in the SP condition coincide with the beginning of sentences, the percentage of false starts made in the SP condition was inflated. Thus, false starts as defined in Experiment 1 conflates two separate phenomena seen in this experiment: rereading triggered by a problem with the word at the beginning of a line, and rereading that word triggered by a problem encountered some words further along. We averted the confound created by this error correction strategy by slightly altering our criteria for false starts and other dysfluencies thereby enabling us to capture the phenomenon of interest.

Accordingly, false starts were redefined as hesitations or stumbles on the first word of a line, or rereading the end of the previous line. If a participant stumbled two or more words into a sentence and then returned to the beginning of the sentence, this sentence onset return was not counted as a false start. Using the revised criteria, a significant and large effect of format was obtained for false starts, $t(22) = 3.25$, $p < .01$, $d = .75$. Participants made more than twice as many false starts in the PD condition (54) than the SP condition (24).[5]

As in Experiment 1, other dysfluencies were defined as hesitations within a line, or stumbles on or rereading any word other than the first word of a line. To be consistent with the new criteria for false starts, sentence onset returns occurring in the middle of a line were not scored as other dysfluencies. There was no difference in percentage of other dysfluencies in the two conditions, $t(22) = .82$, $p = .42$, $d = .17$. There was also no significant effect of format on the comprehension measure, in this case recall ratings, $t(22) = .59$, $p = .56$, $d = .14$.

*Correlations.* Table 4 shows correlations among the experimental measures (WCPM, word errors, fluency ratings, false starts, other dysfluencies, recall ratings) and skill measures (vocabulary, spelling, Gates–MacGinitie comprehension, word attack). The experimental measures were averaged across format.

There are numerous significant correlations among the experimental measures. As in Experiment 1, WCPM correlated significantly with word errors, fluency ratings, false starts, and other dysfluencies. In addition, word errors and fluency ratings each correlated significantly with false starts and other dysfluencies. False starts were significantly correlated with other dysfluencies. Thus, there is a substantial amount of interrelatedness among the fluency measures. Moreover, the skill measures correlated with the fluency measures to a greater extent than in Experiment 1. Each skill measure correlated significantly with WCPM, word errors, fluency ratings, and other dysfluencies. Gates–MacGinitie comprehension and word attack also correlated significantly with false starts. However, the experimental comprehension measure, recall ratings, failed to correlate with any skill or fluency measures.

## Discussion

Consistent with Experiment 1, the results of Experiment 2 demonstrate that demarcating major syntactic boundaries with line breaks had a favorable impact on oral reading fluency. When phrasal constituents were interrupted by a line break, participants obtained lower fluency ratings and made more than twice as many false starts compared to syntactically formatted text. In support of Schreiber's (1991) contention, supplying even a modicum of visual support for parsing by adjusting ends of lines to coincide with ends of phrases seems to facilitate smooth reading of text. However, both Experiments 1 and 2 showed that the text manipulation did not have a consistent impact on measures of reading speed (WCPM) or accuracy (word errors). Instead, it affected aspects of phrasing and prosody, captured by fluency ratings and false starts.

Although the text manipulation had varying effects on different aspects of fluency, overall there were numerous significant correlations among the experimental fluency measures (WCPM, word errors, fluency ratings, false starts, other dysfluencies). Clearly, these separate aspects of fluency are related, supporting the idea that there is a unity underlying its various dimensions. The fluency measure showing the strongest relationship with other fluency measures and with skill measures was WCPM. It appears to capture key elements of fluency and reading-related skills in confirmation of earlier studies (e.g., Fuchs et al., 2001; Torgesen et al., 2001; Young & Bowers, 1995). A comparison of the correlation patterns found in Experiments 1 and 2 shows consistently stronger associations among fluency and skill measures in Experiment 2. This suggests that fluency and other reading-related measures appear to converge in older, more skilled readers to a greater extent than in younger, less skilled readers.

## GENERAL DISCUSSION

This research was designed to explore the possibility that the strategic placement of line breaks on the printed page would promote fluency in reading aloud by

Table 4. *Correlations among fluency, comprehension, and skill measures in Experiment 2*

| Measures | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. WCPM | — | −.62** | .90** | −.72** | −.82** | .18 | .66** | .71** | .74** | .79** |
| 2. Word errors | | — | −.41 | .46* | .53* | −.05 | −.44* | −.72** | −.67** | −.66** |
| 3. Fluency ratings | | | — | −.67** | −.70** | .18 | .65** | .60** | .66** | .68** |
| 4. False starts | | | | — | .79** | −.16 | −.39 | −.43 | −.53** | −.59** |
| 5. Other dysfluencies | | | | | — | −.10 | −.45* | −.56** | −.54** | −.61** |
| 6. Recall rating | | | | | | — | .20 | .01 | .06 | .14 |
| 7. Vocabulary | | | | | | | — | .54** | .48* | .51* |
| 8. Spelling | | | | | | | | — | .52* | .82** |
| 9. Gates–MacGinitie comp. | | | | | | | | | — | .61** |
| 10. Word attack | | | | | | | | | | — |

*$p < .05$. **$p < .01$.

developing readers. To this end, children's text was presented in two formats: one in which major syntactic boundaries coincided with line breaks (SP condition), and the other in which ends of lines interrupted a clausal or phrasal unit (PD condition). The principal result of Experiment 1 was replicated in Experiment 2: there were higher fluency ratings and markedly fewer dysfluencies at line breaks (false starts) when text appeared in the SP condition than the PD condition. Supporting syntactic structure by placement of line breaks resulted in more phrasal readings of text, which included fluid transitions from the end of one line to the beginning of the next. These findings proved to be quite robust; they were observed at two grade levels with materials that differed in length and complexity.

Frequency of dysfluencies occurring elsewhere (i.e., not at the beginnings of lines) did not differ across formats in either experiment. In view of this, we can rule out the possibility that false starts occurred more often in the PD condition simply because difficult words or phrases happened to occur with greater frequency at the beginnings of lines in this condition. If this were the case, then we would anticipate more "other dysfluencies" in the SP condition, where words and phrases found at the beginnings of lines in the PD condition occur *within* the lines in the SP condition. However, such a difference did not occur, indicating that false starts are specifically associated with movement from line to line, and not with difficulties tied to particular words at line breaks.

It should be noted, however, that false starts and other dysfluencies showed less internal consistency than other experimental measures (e.g., WCPM, fluency ratings). Contributing to the instability of these measures, undoubtedly, is the fact that each child read only two passages in each condition. Not only do the passages vary somewhat in the tendency to evoke false starts and other dysfluencies, as would be expected, but a given passage (perhaps based on the inclusion of particular words or idiomatic expressions) seemed to create more difficulty for some children than for others. Differences associated with the particular passages used would certainly be expected. The basis of these idiosyncratic effects will have to be investigated further.

Marking off clauses by the placement of line breaks is a seemingly inconspicuous means of syntactic cuing, yet this manipulation is shown to have a reliable impact on fluent reading. In this connection, eye-movement studies have shown that there is a wrapup effect at a clause or sentence boundary: readers tend to make a longer fixation on a word when it is at the end of a clause than when the same word occurs elsewhere, a result that is generally interpreted to reflect cognitive integration time in completing a parse (Rayner, Kambe, & Duffy, 2000). In addition, the movement of the eyes from the end of one line to the beginning of the next is a complex maneuver that often results in an undershoot on the return sweep, such that the first word of a new line may be skipped, sometimes requiring the reader to make a corrective movement (Rayner, 1998). These two phenomena may have cumulative effects on parsing operations during reading. One effect could surely be the phenomenon we observed: hesitations and repetitions at the beginning of a line that occur more frequently when the line break interrupts a clause, thereby delaying integration until after the return sweep. In contrast, arranging the text to allow the reader to close off a major linguistic constituent before proceeding to the next line reduces false starts and results in higher fluency ratings, arguably

because processing of a major syntactic unit is completed before the reader has to undertake the costly maneuver of executing (and perhaps correcting) a return sweep. These speculations are consistent with our observations and available data on eye movements in reading. Confirmation will have to await further research that systematically varies sentence structure at the end of line break points.

Although clear format effects were obtained for fluency ratings and false starts, no consistent effects of format were found for WCPM.[6] Similarly, no significant effect of format was found for word errors in either experiment. These findings suggest that reading fluency may consist of somewhat separable components: (a) speed and accuracy of word reading (measured by WCPM and word errors) and (b) phrasing and prosody (captured by fluency ratings and false starts). Our findings are consistent with commonly accepted definitions of reading fluency (see National Reading Panel, 2000), which stress that fluency entails both reading accuracy/speed as well as proper expression.

Although we have evidence for distinct aspects of fluency, we also find in our correlational analyses that children tend to show overlapping skills in these two areas of fluency. Significant correlations were found among the experimental measures of fluency in each experiment, and these correlations were higher for the older children in Experiment 2. Our findings suggest that the various aspects of fluency emerge separately in younger children and converge later on. Thus, for example, a less skilled reader may read relatively fast, but with poor phrasing or unnatural prosody. Thus, the use of syntactically formatted materials may afford a benefit to continuity and phrasing while reading aloud, independently of reading rate. We conjecture that in older, poor readers we may continue to see discrepancies between phrasing and reading speed.

In addition to finding that the experimental measures of fluency showed higher intercorrelations in Experiment 2 than Experiment 1, we also found that correlations between the experimental measures and language and reading skill measures (vocabulary, spelling, Gates–MacGinitie comprehension, word attack) were somewhat higher in Experiment 2. In particular, WCPM showed strong and consistent relationships with all skill measures, confirming its value as a summary index of fluency (see Fuchs et al., 2001; Torgesen et al., 2001; Young & Bowers, 1995). Likewise, fluency ratings were significantly correlated with all skill measures in Experiment 2. Thus, we found evidence that both components of fluency (speed/accuracy and phrasal reading) are related to general reading and language skills in older children.[7]

Although oral reading fluency was the chief focus of our experiments, of course we were interested to learn whether syntactic formatting confers benefits to reading comprehension as well. It seems compelling that reading with appropriate phrasing and prosody is evidence that at least syntactic analysis (phrase structure) has occurred. Further, appropriate placement of contrastive stress is evidence that pragmatic features of focus and emphasis have been apprehended by the reader. Both are central for comprehension by any linguistic account of which we are aware (see Fodor, 2002). However, we found no discernible effects of format on our comprehension task, either when the task was to answer specific questions (Experiment 1) or to recall the story (Experiment 2).

It is likely that most of the children in our experiments were functioning far enough above the level of our text materials that they were able to recover and repair

any disruption of comprehension introduced by end of line breaks, especially in Experiment 2. In both experiments, average performance on the Gates–MacGinitie comprehension subtest and vocabulary was above grade level. It seems likely that children at earlier stages of development would be more adversely affected by end of line breaks, both in terms of reading rate and comprehension measures. Another likely possibility is that our students would show greater effects of text manipulation on comprehension if asked to read under time pressure, such that rereading and offline reflection are minimized. In fact, there is evidence that WCPM correlates best with comprehension when readers are under time constraints (Blommers & Lindquist, 1944).

We acknowledge that the present findings on the effects of text end of line formatting on oral reading fluency are preliminary, and that a greater variety of readers is needed to determine who could most benefit from the manipulation. Perhaps there already exists enough evidence to suggest that there may be benefits of broadening the classroom use of line-formatted text to facilitate the transition from unconnected word by word reading to fluent, phrasal reading. Although the most elementary reading materials (intended for first grade) often present one sentence per line, second-grade texts are usually produced without regard to linguistic structure. When examining our materials in their unaltered, published form, we found that only 31% of the lines ended in intact syntactic units, making these texts more similar to our PD condition than our SP condition. In view of our findings, it may be wise to delay the transition to unformatted text until children have demonstrated adequate reading rates and show evidence of phrasal reading.

The utility of practice with phrase-cued text is also potentially relevant to the classroom. Repeated work with syntactically formatted text may afford transitional readers greater benefits than the limited exposure in the present study. To investigate this possibility, the first author conducted an intervention study that combined aspects of text formatting (end of line cuing to clausal structure and overt cuing of phrasal structure) with repeated readings to determine whether practice with these materials would benefit fluency and show transfer to conventional (unformatted) text. It was found that practice with formatted text led to more fluent reading of conventional text after training. Fluency ratings and false starts each benefited significantly more from practice with segmented text than practice with ordinary text (LeVasseur, 2004). Further studies are planned that focus on individual differences.

## APPENDIX A

### PASSAGE AND COMPREHENSION QUESTIONS USED IN EXPERIMENT 1

The following is one of the passages used in the formats, shown in SP and PD. The comprehension questions follow. The other passages used in the experiment are available upon request.

*Passage 1* in SP format:

Dear Oscar and Tammy,
I miss you both a lot.

After our phone call
I really felt like packing my suitcase
and going to join you at Grandpa and Grandma's house.
I know that you would have rather stayed at home
this vacation, but being alone is just too dangerous.
It's safer for you to be at Grandpa's house.
You will have fun.
One more thing—
it is too expensive to call on the phone.
I am sending you some writing paper, envelopes,
and stamps.
I am also sending some new crayons and drawing paper for Tammy.
Write soon!

<div align="right">Love, Mom</div>

*Passage 1* in PD format:

Dear Oscar and Tammy,
I miss you both a
lot. After our phone call I really felt
like packing my suitcase and
going to join
you at Grandpa and Grandma's
house. I know that you would have rather stayed at
home this vacation, but being alone
is just too dangerous. It's safer for you to
be at Grandpa's house. You will have fun. One more
thing—it is
too expensive to call on the phone. I
am sending you some writing
paper, envelopes, and stamps. I am
also sending some new crayons and drawing
paper for Tammy. Write soon!

<div align="right">Love, Mom</div>

Practice Paragraph:

1. How are Oscar and Tammy related?
2. What will they be doing while they are away for the summer?

Passage 1:

1. Where are Oscar and Tammy staying?
2. Why did Oscar and Tammy go to Grandpa and Grandma's house?
3. Why didn't Mom want Oscar and Tammy calling on the phone?
4. What did Mom send to the children?

Passage 2:

1. How did Oscar and Tammy help Grandma?
2. Who showed Oscar and Tammy how to make vegetable soup?
3. What did Tammy see in the dirt?
4. What happened to Tammy when Grandma told her to taste the soup?

Passage 3:

1. When did Grandpa make his soup?
2. When did Mom work in the vegetable garden?
3. What gift did Mom receive at the party?
4. Why does Mom's house seem very quiet?

Passage 4:

1. What was Grandpa trying to paint?
2. How did Tammy get in trouble?
3. What color was the paint?
4. Who made a drawing for Mom?

## APPENDIX B

## AN EXAMPLE OF A PASSAGE USED IN EXPERIMENT 2

The following excerpt from *The Midnight Pig* is shown first in the SP and next in the PD format. Copies of the other passages used in the experiment are available upon request.

*The Midnight Pig* in SP format:

We called Fred the "Midnight Pig"
because that's when he usually ate. We owned the local store,
and by the time Mom had closed up
and prepared for the next day, it was usually about midnight.
The old clock would strike twelve,
and Mom would throw out cabbage leaves, vegetable peels, and any other
leftover food. Fred
would wolf it down.

I can still remember the day I got Fred.
It was my first day at my new school.
We had just moved from the city.
Although Mom was hesitant at first,
she finally agreed to me having a pet pig.

As time went by, things started to settle down,
and I became quite happy in our new town.
Fred, however, had a habit of making life difficult for me.

I got a part in the school play,
so I went to rehearsal after school.
The play went really well—sort of.
We performed at the Town Hall,
in front of all the parents.
The only thing that ruined it was Fred.

He must have followed Mom to the Town Hall and stood,
watching through the open doorway.
In the final scene, I had to collapse to the floor.
Fred was obviously concerned.
He clip-clopped across the floor and onto the stage.
He kept grunting and butting my face with his wet nose.
Everyone laughed and applauded.
Mark had to pull him off the stage.
Boy, was I embarrassed!

*The Midnight Pig* in PD format:

We called Fred the "Midnight
Pig" because that's when he usually
ate. We owned the local store, and by the time
Mom had closed up and prepared for the next
day, it was usually about midnight. The old
clock would strike twelve, and Mom would throw
out cabbage leaves, vegetable
peels, and any other leftover
food. Fred would wolf it down.

I can still remember the day I got
Fred. It was my first day at my new school. We had just
moved from the city. Although Mom was hesitant at
first, she finally agreed to me having a pet pig.

As time went by, things started to settle
down, and I became quite happy in our
new town. Fred, however, had a habit of making
life difficult for me. I got a
part in the school play, so I went to rehearsal after
school. The play went really well—sort of. We performed at
the Town Hall, in front of all the parents. The only
thing that ruined it was Fred.

He must have followed Mom to the Town Hall and
stood, watching through the open
doorway. In the final scene, I had to
collapse to the floor. Fred was obviously
concerned. He clip-clopped across the
floor and onto the stage. He kept grunting and

butting my face with his wet nose. Everyone
laughed and applauded. Mark had to pull
him off the stage. Boy, was I embarrassed!

NOTES
1. Normed data are not available for the DSA.
2. Correlations were also computed for the SP and PD conditions separately. The patterns
   of correlations did not differ enough from those obtained with combined conditions to
   warrant reporting them separately.
3. A minor modification was made to *The Week of the Jellyhopper* (Cartwright, 1995)
   Two city names in the original text that are unfamiliar (Opito, Mason's Flat) were
   replaced with names more familiar to the children (Newport, Cape Cod).
4. Forty-seven percent of false starts in Experiment 2 were sentence onset returns com-
   pared to only 17% for Experiment 1. Sentence onset returns did not differ significantly
   across format. In Experiment 2, there were on average 2.0 ($SD = 1.5$) sentence onset
   returns in the SP condition and 2.3 ($SD = 2.0$) in the PD condition.
5. In both Experiments 1 and 2, the percentage of false starts and word errors showed
   positively skewed distributions (i.e., many were 0%, especially in the SP condition). To
   reduce this skew, we employed a square root transformation on percentage scores. Our
   $t$ tests on transformed scores produced the same pattern of results as untransformed
   data. Results reported in this study are based on untransformed data.
6. In a previous study with adults, we did find a significant effect of formatted text on
   reading rate. In a self-paced reading situation, participants read either intact or disrupted
   phrases as they appeared on a computer screen. Reading rates were significantly faster
   when phrases were kept intact at the ends of lines than when they were disrupted.
   We speculate that one reason for this divergent finding is due to the differences in
   methodology. The adult study allowed for millisecond accuracy in reading times: both
   line by line and overall passage times (Marciarille-LeVasseur et al., 2001).
7. We also examined individual differences in text format effects as they relate to language
   and reading skill measures. Difference scores (PD − SP) were computed on fluency
   measures that showed significant format effects (i.e., false starts and fluency ratings
   in both experiments). The only significant correlation was between vocabulary scores
   and PD − SP difference in false starts in Experiment 2 ($r = .49$, $p < .01$). Children
   with lower vocabulary scores showed the largest reduction in false starts in the SP
   condition. This finding requires further investigation for a clearer interpretation.

# REFERENCES

Almada, P. (1997). *Oscar and Tatiana*. Crystal Lake, IL: Rigby Education.

Bever, T. G., Jandreau, S., Burwell, R., Kaplan, R., & Zaenen, A. (1990). Spacing printed text to isolate major phrases improves readability. *Visible Language, 25*, 74–87.

Blommers, P., & Lindquist, E. F. (1944). Rate of comprehension of reading: Its measurement and its relation to comprehension. *Journal of Educational Psychology, 35*, 449–473.

Cartwright, P. (1995). *The week of the jellyhopper*. Auckland, New Zealand: Shortland Publications.

Carver, R. P. (1970). Effect of a "chunked" typography on reading rate and comprehension. *Journal of Applied Psychology, 54*, 288–296.

Chall, J. S. (1983). *Stages of reading development*. New York: McGraw–Hill.

Chomsky, C. (1976). After decoding: What? *Language Arts, 53*, 288–296.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Critchlow, D. E. (1996). *Dos Amigos Verbal Language Scales*. Novato, CA: Academic Therapy Publications.

Dowhower, S. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Research Quarterly, 22*, 389–406.

Dugan, M. (1988). *It's just a trick*. Crystal Lake, IL: Rigby Education.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., et al. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86*, 6–23.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. In M. Hirotani (Ed.), *Proceedings of NELS 32*. Amherst, MA: University of Massachusetts, GLSA.

Foorman, B. R., Francis, D. J., Novy, D., & Liberman, D. (1991). How letter-sound instruction mediates progress in first-grade reading and spelling. *Journal of Educational Psychology, 83*, 456–469.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*, 239–256.

Fuchs, L. S., Fuchs, D., & Maxwell. (1988). The validity of informal measures of reading comprehension. *Remedial and Special Education, 9*, 20–28.

Ganske, K. (1999). The Developmental Spelling Analysis: A measure of orthographic knowledge. *Educational Assessment, 6*, 41–70.

Gough, P. B. (1985). One second of reading. In H. Singer & R. B. Ruddel (Eds.), *Theoretical models and processes of reading* (3rd ed., pp. 661–686). Newark, DE: International Reading Association.

Graf, R., & Torrey, J. W. (1966). Perception of phrase structure in written language. *Proceedings of the Annual Convention of the American Psychological Association, 74*, 83–84.

Jandreau, S., & Bever, T. (1992). Phrase-spaced formats improve comprehension in average readers. *Journal of Applied Psychology, 77*, 143–146.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon.

Klein, R. (1995). *Glumly. USA*. Auckland, New Zealand: Mimosa Publications.

Krueger, C. (1997). *The midnight pig*. Crystal Lake, IL: Rigby Publications.

Laberge, D., & Samuels, S. J. (1985). Toward a theory of automatic information processing in reading. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd ed., pp. 689–718). Newark, DE: International Reading Association.

LeVasseur, V. M. M. (2004). *Promoting gains in reading fluency by rereading phrasally cued text*. Unpublished doctoral dissertation, University of Connecticut.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L.G. (2000). *Gates–MacGinitie Reading Tests* (4th ed.). Itasca, IL: Riverside Publishing.

Marciarille-LeVasseur, V. M., Shankweiler, D., & Macaruso, P. (2001, April). *Piecemeal reading*. Paper presented at the meeting of the Eastern Psychological Association, Washington, DC.

Mason, J. M., & Kendall, J. R. (1979). Facilitating comprehension through text structure manipulation. *The Alberta Journal of Educational Research, 25*, 68–76.

Meyer, M. S., & Felton, R. (1999). Repeated reading to enhance fluency: Old approaches and new directions. *Annals of Dyslexia, 49*, 283–306.

National Center for Education Statistics. (1995). *Listening to children read aloud*. Retrieved August 7, 2003, from http://nces.ed.gov/pubs95/web/95762.asp#tab1.

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research on reading and its implications for reading instruction.* Washington, DC: National Institute of Child Health and Human Development.

Odgers, S. F. (1987). *What a day!* Auckland, New Zealand: Shortland Publications.

O'Shea, L. J., & Sindelar, P. T. (1983). The effects of segmenting written discourse on the reading comprehension of low- and high-performance readers. *Reading Research Quarterly 18*, 458–465.

Rasinski, T. V. (1994). Developing syntactic sensitivity in reading through phrase-cued texts. *Intervention in School and Clinic, 29*, 94–97.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology, 53A*, 1061–1080.

Rumelhart, D. E. (1985). Toward an interactive model of reading. In H. Singer & R. Ruddell (Eds.), *Theoretical models and processes of reading* (3rd ed., pp. 722–750). Newark, DE: International Reading Association.

Schnick, T., & Knickelbine, M. (2000). *The lexile framework: An introduction for educators.* Durhan, NC: MetaMetrics, Inc.

Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior, 12*, 177–186.

Schreiber, P. A. (1991). Understanding prosody's role in reading acquisition. *Theory into Practice, 30*, 158–164.

Schreiber, P. A., & Read, C. (1980). Children's use of phonetic cues in spelling, parsing, and—maybe—reading. *Bulletin of the Orton Society, 30*, 209–224.

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., et al. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading, 3*, 69–94.

Stanovich, K. E., Cunningham, A. E., & Feeman, D. J. (1984). Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly, 19*, 278–301.

Stanovich, K. E., & Siegel, L. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of phonological-core variable-difference model. *Journal of Educational Psychology, 86*, 24–53.

Torgesen, J. K., Rashotte, C. A., & Alexander, A. W. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 333–355). Parkton, MD: York Press.

Weiderholt, J. L., & Bryant, B. R. (2001). *Gray Oral Reading Tests, 4th edition.* Austin, TX: PRO-ED, Inc.

Weiss, D. S. (1983). The effects of text segmentation on children's reading comprehension. *Discourse Processes, 6*, 77–89.

Wood, C. T. (1975). Processing units in reading. (Doctoral dissertation, Stanford University). *Dissertation Abstracts International, 35*, 46–99.

Woodcock, R., & Johnson, M. B. (1993). *Woodcock–Johnson Psychoeducational Battery.* Itasca, IL: Riverside Publishing.

Young, A., & Bowers, P. G. (1995). Individual differences and text difficulty determinants of reading fluency and expressiveness. *Journal of Experimental Child Psychology, 60*, 428–454.