Paige, David D.; Rupley, William H.; Smith, Grant S.; Rasinski, Timothy V.; Nichols, William; Magpuri-Lavell, Theresa

Is prosodic reading a strategy for comprehension?

Journal for educational research online 9 (2017) 2, S. 245-275

Journal for Educational Research Online
Journal für Bildungsforschung Online
Volume 9 (2017), No. 2, 245–275

David D. Paige, William H. Rupley, Grant S. Smith, Timothy V. Rasinski, William Nichols & Theresa Magpuri-Lavell

Is prosodic reading a strategy for comprehension?

Abstract

Emerging research on reading prosody, an indicator of fluent reading, is finding that it contributes to comprehension processing in students across elementary, middle, and secondary grades. In this study we measure the growth of reading prosody and comprehension of 250 first-, second-, and third-grade readers across the school year using the Multidimensional Fluency Scale (MDFS; Zutell & Rasinski, 1991). Our results show that students gradually improve their reading prosody and reach asymptote with grade-level text by the end of second-grade. We found that reading rate was not a significant predictor of comprehension while word identification accuracy and prosody accounted for 64.9% of unique variance in reading comprehension. Using both a three-step (Baron & Kenny, 1986) and bootstrap resampling approach to mediation analysis (Preacher &

David D. Paige, Corresponding Author
Bellarmine University, 2001 Newburg Road, Louisville, Kentucky USA 40205
e-mail: dpaige@bellarmine.edu

William H. Rupley
Texas A&M University, 206 Harrington Tower, Education and Human Development, Texas A&M University, College Station, Texas USA 77840-77845
e-mail: w-rupley@tamu.edu

Grant S. Smith
Bellarmine University, 2001 Newburg Road, Louisville, Kentucky USA 40205
e-mail: gssmith@bellarmine.edu

Timothy V. Rasinski
Kent State University, 401 White Hall Kent, Ohio USA 44242
e-mail: trasinsk@kent.edu

William Nichols
University of Maine, Merrill Hall Orono, Maine USA 04469-5766
e-mail: william.nichols1@maine.edu

Theresa Magpuri-Lavell
Bellarmine University, 2001 Newburg Road, Louisville, Kentucky USA 40205
e-mail: tmagpurilavell@bellarmine.edu

Hayes, 2004; Preacher & Kelly, 2011), we found that prosody exerts a significant mediating effect on the relationship between automaticity and comprehension. Additional analysis revealed that a reader's ability to accurately read connected text with appropriate pacing emerges more quickly than does expressive reading and phrasing. Finally, we use the Implicit Prosody Hypothesis (Fodor, 2002) to advance the notion that prosodic readers may leverage reading prosody as a problem-solving tool to interpret ambiguous text, and thus increase reading comprehension.

Keywords
Prosody; Fluency; Comprehension; Elementary grades; Reading comprehension

Prosodisches Lesen als Strategie für das Leseverständnis?

Zusammenfassung
Die aufstrebende Forschung zur Leseprosodie als ein Indikator für flüssiges Lesen zeigt, dass Prosodie bei Schülerinnen und Schülern verschiedener Altersklassen zum Leseverständnis beiträgt. In der vorliegenden Studie wurde unter Verwendung der Multidimensional Fluency Scale (MDFS; Zutell & Rasinski, 1991) die Entwicklung der Leseprosodie und des Leseverständnisses von 250 Erst-, Zweit- und Drittklässlern über ein Schuljahr hinweg untersucht. Die Ergebnisse zeigen, dass die Schülerinnen und Schüler schrittweise ihre Leseprosodie verbessern und sich bis zum Ende der zweiten Klasse dem der Jahrgangsstufe entsprechenden Niveau annähern. Während sich die Lesegeschwindigkeit nicht als signifikanter Prädiktor des Leseverständnisses erwies, konnten die Genauigkeit der Wortidentifikation und die Leseprosodie 64.9 % der Eigenvarianz beim Leseverständnis aufklären. Mithilfe eines Three-Step-Resamplings (Baron & Kenny, 1986) und eines Bootstrap-Resamplings zur Mediationsanalyse (Preacher & Hayes, 2004; Preacher & Kelly, 2011) konnte ein signifikanter Mediationseffekt der Prosodie auf das Verhältnis von Automatizität und Leseverständnis gezeigt werden. Weitergehende Analysen verdeutlichten, dass sich Lesegenauigkeit und angemessenes Lesetempo schneller entwickeln als die Fähigkeiten zu expressivem Lesen und Phrasierung. In Anlehnung an die Implicit Prosody Hypothesis (Fodor, 2002) wird der Gedanke weitergeführt, dass Leseprosodie als problemlösendes Hilfsmittel zum Interpretieren nicht-eindeutiger Wortlaute eingesetzt und das Leseverständnis dadurch verbessert werden kann.

Schlagwörter
Prosodie; Lesefluss; Verständnis; Primarstufe; Leseverständnis

1. Introduction

The expectation of the Common Core State Standards (CCSS; CCSS, 2010) is that students depart 12th-grade reading at a level that is college- and career-ready (CCR; see Appendix A of CCSS, 2010 for more information). Unfortunately, this does not happen for many students (ACT, 2013). To bring coherence and intentionality to achieving CCR, the CCSS has quantified reading attainment using Lexile bands that make explicit the complexity level of texts by grade level. A Lexile is a quantitative metric representing 1/1000th of the difference between the mean difficulty of primer reading material and that of an encyclopedia (MetaMetrics, 2012). Within the foundational skills (K-5) of the standards is the expectation that students will fluently read grade-level text. While much is known about the phonological development of early readers that contributes to accuracy and automaticity (see Adams, 1990; Dickinson & Neuman, 2006; Stanovich, 2000; Torgesen & Hecht, 1996), less is known about the development of prosodic reading in early readers and its possible relationship with other fluency indicators and comprehension. This study explores the development of reading prosody and its effect on comprehension in first-, second-, and third-grade students.

2. Theoretical background

2.1 Decoding, Automaticity, and Fluency

Accurate and automatic decoding are fundamental capabilities for fluent reading because they allow the reader to focus their attention on text comprehension (Logan, 1988; Schneider, 1985). It is visual word identification that is the "most distinctive process for reading" (Perfetti, 1999, p. 170). In a model described by Perfetti (1999), successful word identification begins with the visual loading of a string of letters that then activates word forms in the mental lexicon. The reader develops accuracy with word identification through intentional and sequential phonics instruction where they learn to interpret consistencies at both the phoneme and rime level (Goswami, 2005). Awareness of phoneme and rime consistencies results in the ability to string together letters into recognizable units which provides the reader access to the lexicons of pronunciation and meaning (Compton, Appleton, & Hosp, 2004; Van Orden & Goldinger, 1994). However, if the process ended here, short term mental resources would be heavily invested in word identification and comprehension would be minimized. Thus, automaticity must develop to relieve overburdening the mental processes at the pronunciation and individual word meaning level.

Automaticity is the ability to perform a complex skill with little conscious mental attention and effort (Samuels & Flor, 1997). Logan (1988) submits that automatic processing is a) rapid activation, b) effortless and autonomous, and c) done without conscious intention. Word automaticity is a reader's ability to instantaneously retrieve a word from long-term memory as a whole unit, in other words, the word becomes available to the reader before a decoding algorithm is applied to unlock the word from print (Logan, 1988). Automatic processes, and specifically word identification automaticity, are developed through practice (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977).

Huey (1968), investigating reading in the early 20th century, recognized that fluency "frees the mind from attention to details, makes facile the total act, shortens the time, and reduced the extent to which consciousness must concern itself with the process" (p. 104). Embedded in this statement is the notion that increasingly efficient decoding nurtures automaticity at the word level that in turn, enhances reading fluency which allows the reader to devote more cognitive attention to comprehension. As such, it frees the mind to engage in comprehension experiences (Heilman, Blair, & Rupley, 2002). Thus, a virtuous cycle is created where increased word automaticity facilitates greater reading fluency that frees attentional resources to focus mental resources on meaning (LaBerge & Samuels, 1974; Logan, 1988; Logan & Etherton, 1994; Samuels, 2006; Torgesen & Hudson, 2006). Because of its relationship to comprehension the National Reading Panel (2000) identified reading fluency as fundamental to effective reading instruction (Daane, Campbell, Grigg, Goodman, & Oranje, 2005; Fuchs, Fuchs, Hosp, & Jenkins, 2001; Miller & Schwanenflugel, 2006; Paige, 2011; Paige, Rasinski, Magpuri-Lavell, & Smith, 2014; Pinnell et al., 1995; Schatschneider et al., 2004; Stecker, Roser, & Martinez, 1998; Young, Bowers, & MacKinnon, 1996).

Developing in parallel with a focus on reading fluency has been an evolving understanding of what it is. For example, a measurement view of fluency reflects a reader's accuracy in decoding words across the time taken to read connected text (Hasbrouck & Tindal, 2006; Torgesen, Rashotte, & Alexander, 2001). The *rate concept* of fluency ignores prosody which is the ability to reflect conversational language when reading (Rasinski, Reutzel, Chard & Linan-Thompson, 2011). Breznitz (2006) moves beyond reading rate and takes a broad perspective of fluency that considers a multitude of cognitive and linguistic sub-processes that work in synchronization. This perspective takes into account readers' efficiency in using lexical features of text (Perfetti, 1985, 1988). Many researchers today concur that the construct of reading fluency is manifested in the indicators of automaticity (rate), word identification accuracy, and prosody, and acknowledge its importance to comprehension (Hudson, Pullen, Lane, & Torgesen, 2009; Kuhn & Stahl, 2003; Paige et al., 2014; Samuels, 2007; Schwanenflugel et al., 2006).

2.2 Prosody: Reading fluency's missing factor

A simplistic view of prosody is *reading with feeling*. Perhaps for competent readers this is a succinct and adequate definition while others may think of prosody as the "ability to make oral reading sound like authentic oral speech" (Rasinski et al., 2011, p. 292). Dowhower (1991) defines prosodic oral reading as recognition of markers identifying appropriate use of pauses, phrasing, pitch, and stress. Speech production and communication of meaning are built on these components (Vihman, 2014). Finding answers about how competent readers get to this point can lead to advancing our understanding on its acquisition and contribution to fluent reading.

Prosody development and utilization in understanding oral language begins at birth. Initial oral language struggles encountered by an infant are the isolation and identification of spoken words from the ambient environment. Speech has no easily discernible segmentation which makes oral word identification a challenging task, but one the infant begins to untangle almost immediately after birth (Cole & Jakimik, 1980). By three to four days of age, newborns prefer their native prosodic pattern to those that are foreign (Christopher, Mehler, & Sebastian-Galles, 2001; Mehler et al., 1988). Infants living in an English-language environment quickly learn that approximately 90 % of the time words are stressed on the initial syllable (strong-weak), a helpful insight that allows them to use lexical stress to break the stream of speech in their aural environment and enhances its understanding (Cutler & Carter, 1987; Thiessen & Saffran, 2003). Sensitivity to stress represents a prosodic template on to which infants can fit strings of speech (Vihman & Vellman, 2000). These strings become prosodic phrases that enable faster encoding into memory (Mandel, Jusczyk, & Nelson, 1994; Mandel, Jusczyk, & Pisoni, 1995). The months between five and nine are marked by significant language learning as infants learn to use lexical stress to discriminate between strong and weakly stressed syllables (Weber, Hahne, Friedrich, & Friederici, 2004). Lexical stress discrimination informs conditional or transitional probability (when given one event a second is likely to occur) to assist the infant in learning to identify word boundaries (Aslin, Saffran, & Newport, 1998). Also during this short time period infants learn to identify specific words (Jusczyk & Aslin, 1995) and distinguish between them based on the probability that many syllables occur together (Aslin, et al., 1998). Prosody development in speech for infants culminates at approximately nine months when they have acquired significant knowledge of the prosodic properties of their respective language that is hypothesized to be important to fluent speech and communication (Church, 1987; Cutler & Butterfield, 1992). Further, those children exposed to English prefer to hear words with a strong/weak stress, what are known as trochees, as opposed to a weak/strong stress or iambs (Jusczyk, Cutler, & Redanz, 1993). Even at this early age, children who are read to from storybooks and nursery rhymes are remembering some of the words and patterns of language, thus beginning the critical development of vocabulary and syntax (Jusczyk & Hohne, 1997). In sum, infants learn to use acoustical cues to disambiguate words from the stream of speech occurring in their environment, eventually learning that mutually predictive syllables are likely to co-occur or cluster together. Before one year of age, the child has learned to use prosodic markers to untangle oral speech; markers that will later be applied to reading of text. Increasing evidence reveals that prosodic sensitivity found in early speech is important to phonological awareness, and eventually to accuracy in word reading and spelling (Holliman et al., 2014). For example, Wade-Wooley (2016) found that after controlling for general word reading ability, children who could correctly identify the placement of lexical stress were better able to read multisyllabic words. While prosody is often the neglected component of reading fluency, it has always been connected to language and effective communication. Prosody, specifically the rhythm of speech, has been found to predict unique variance in word reading and the ability to phrase connected text, both important indicators of fluent reading (Dowhower, 1991; Holliman, Wood, & Sheehy, 2010; Miller & Schwanenflugel, 2008; Zutell & Rasinski, 1991).

Prosodic reading occurs when the reader applies characteristics of prosodic speech to both oral and silent reading (Bader, 1998; Carlson, 2009; Frazier, Carlson, & Clifton, 2006; Rasinski et al., 2011). But what is it then that distinguishes a prosodic reader from one who is not? While reading words accurately and with automaticity is important, many more cognitive processes are involved in prosodic reading (Kuhn & Stahl, 2003; LaBerge & Samuels, 1974; Torgeson, Rashotte, & Alexander, 2001). As with speech prosody, the reader must learn to apply stress to specific syllables, a marker of skilled reading (Himmelmann & Ladd, 2008; Goswami et al., 2002; Whalley & Hansen, 2006; Wood, 2006). Prosodic readers are astute at recognizing informational units in text that can be bracketed into phrases (Frazier et al., 2006; Goldman, Meyerson, & Cote, 2006; Koriat, Greenberg, & Krenier, 2002; Kuhn, Schwanenflugel, & Meisinger, 2010; Schreiber, 1980, 1987, 1991; Swets, Desmet, Hambrick, & Ferreira, 2007; Schwanenflugel & Benjamin, 2016; Schwanenflugel, Westmoreland, & Benjamin, 2015). Bracketing mimics familiar boundaries of oral language by creating a cognitive architecture that aids understanding by parsing text into comprehensible syntactic units (Beckman, 1996; Cutler, Dahan, & van Donselaar, 1997; Peppé & McCann, 2003; Sanderman & Collier, 1997). The ability to phrase and bracket text is thought to be an automatic process that is activated while reading, providing a rhythmic quality to speech that is evidence of normal reading development (Guitérrez-Palma, & Palma-Reyes, 2008). For prosodic reading to occur, information at the word, phrase, and sentence level must be integrated by the reader with semantic information (Kinsch, 1988, 1998). Readers will often pause at the end of a clause or sentence to facilitate this processing (Hirotani, Frazier, & Rayner, 2006; Rayner, Kambe, & Duffy, 2000; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989), offering evidence that whether reading aloud or silently, prosody is critical for comprehension (Benjamin & Schwanenflugel, 2010; Daane et al., 2005; Gross, Millett, Bartek, Bredell, & Winegard, 2014; Pinnell et al., 1995).

Laver (1997) identifies two distinct applications of speech prosody. First, the rise and fall in speaking patterns can serve as a paralinguistic function to indicate humor or irony. A second distinction is the various types of linguistic uses where the speaker conveys meaning through the emphasis of a specific word. One example of this is found in an annoyed request of a waiter to “PLEASE bring me a glass of water.” Inflection at the end of a sentence to indicate a question and the use of a lower pitch to indicate a statement are other examples of linguistic uses (Peppé et al., 2010). This suggests that prosody can be intentionally manipulated by the reader to achieve an identified, desired goal reflecting the emotion of the speaker. While this is a very specific use of prosody, what do we know about the larger usefulness of prosody to reading comprehension?

2.3 Connecting prosody to comprehension

In a survey of 4th-grade readers using a 4-level prosody scale, both Pinnell et al. (1995) and Daane, Campbell, Grigg, Goodman, and Oranje (2005) found that as part of fluent reading, phrasing and oral expression were positively related to comprehension. In a study of 5th-grade students, Klauda and Guthrie (2008) also found that phrasing predicted unique variance in reading comprehension. Valencia et al. (2010) studied 2nd-, 4th-, and 6th-grade students in a cross-sectional study that assessed the power of rate, accuracy, and prosody to predict silent reading comprehension. The authors found that prosody explained significant variance in comprehension that increased between grades, while rate was a non-significant predictor at each of the three grade levels. Veenendaal, Groen, and Verhoeven, (2014) studied 106, normally-achieving 4th-grade Dutch children using the Multi-Dimensional Fluency Scale (MDFS, Zutell & Rasinski, 1991) to measure prosody. Results showed this group to possess very good prosodic reading skills and after controlling for decoding which was a non-significant predictor, vocabulary, syntactic awareness, and prosody explained 51 % of the variance in reading comprehension. Of these three explanatory variables, phrasing uniquely explained 6 % of the total variance. In a second study, Veenendaal, Groen, and Verhoeven (2015) used the summative score from the MDFS and found that combined with vocabulary and syntactic ability, it accounted for 46 % of variance in comprehension, with prosody explaining 3 % of the total variance. Also using the MDFS, Rasinski, Rikli, and Johnston (2009) measured the fluent reading skills of 3rd-, 5th-, and 7th-grade students and found moderately strong correlations to silent reading comprehension. Fluency and comprehension processes of generally struggling 9th-grade readers were studied by Paige et al. (2014) with results showing that reading prosody (measured by the MDFS) and word identification accuracy predicted 52.7 % of the variance in a standardized test of reading comprehension (automaticity was not a significant predictor). After controlling for word identification accuracy, prosody predicted 7 % of unique variance in comprehension, similar to the amount found by Veenendaal et al. (2014) in 4th-grade students. Benjamin and Schwanenflugel (2010) found that after controlling for reading automaticity and word identification accuracy, prosody accounted for an additional 5.5 % of the variance in reading comprehension when 2nd-grades students read an above grade level passage. Interestingly, when reading a grade-level passage prosody explained no variance in comprehension. Recently, Schwanenflugel and Benjamin (2017) found in a study of 3rd-grade students that not only was lexical prosody associated with fluent reading, it also predicted reading comprehension beyond that predicted by rate and accuracy.

In struggling readers poor sensitivity to the rhythmic features of text has been suggested to inhibit development of word-level reading skills (Goswami et al., 2002; Richardson, Thomson, Scott, & Goswami, 2004). Dowhower (1987) found that after students had engaged in repeated reading practice, sensitivity to prosodic features improved through a reduction in inappropriate pauses and an increase in appropriate phrasing and use of pitch (expression). Likewise, Schwanenflugel, Hamilton, Kuhn, Wisenbaker, and Stahl (2004) also found that as decoding skills increased, prosodic reading improved. These results suggest that as the reader improves in their ability to accurately read connected text with automaticity, improvements in prosodic reading may occur. This is consistent with Perfetti's verbal efficiency theory (1985) as well as LaBerge and Samuels (1974) theory of automaticity that both hypothesize that as text processing becomes increasingly automatic, cognitive resources are freed which can then be directed to the prosodic processing of text.

Using the three indicators of fluent reading the tandem theory (Paige et al., 2014) provides an explanation regarding the role of prosody in comprehension. The theory assumes the reader's goal is to a) understand the text and b) invoke appropriate comprehension monitoring while reading. With these in place the reader accurately decodes the words and applies appropriate prosody to the text. This frees the reader to intentionally adjust reading rate up or down in tandem with their perceived level of comprehension. If comprehension occurs easily, the reader may increase their reading rate. However, if comprehension is difficult the rate may be decreased, even to the point of re-reading. Within the theory prosody can be applied by the reader at potentially two levels, first as a foundational cognitive architecture supporting text-base comprehension (Kintsch, 1988) and secondly, as a potential problem-solving tool for resolving textual ambiguity (Fodor, 2002).

During silent reading readers generate mental prosodic-phonological written text representations that Fodor (2002) refers to as implicit prosody (Ashby, 2006; Ashby & Clifton, 2005; Chafe, 1988; Frost, 1998; Savill et al., 2011). While sentence structure is represented by phonological encoding, it is often insufficient in the resolution of ambiguous word structures (Fodor, 2002; Slowiaczek & Clifton, 1980). The implicit prosody hypothesis (IPH; Fodor, 2002) suggests that when reading silently the reader projects a default prosodic contour onto the text that is language specific. However, when faced with sentence ambiguity, the default contour may be inadequate in understanding the text. In such a case the reader may intentionally apply a variety of contours, eventually favoring the one that is most helpful in interpreting the sentence structure. In other words, reader auditions different prosodic interpretations of the text until one is found that seems to be a best fit for meaning. This essentially becomes a problem-solving strategy that provides the reader with a mechanism to better understand the text. It would also seem that the reader must be proficient enough in decoding to actively monitor their comprehension in order to recognize sentence ambiguity in need of resolution. The

IPH suggests that even with little textual cuing, the reader projects prosodic elements onto a text that helps in the resolution of text ambiguity that ultimately assists comprehension. This is hypothesized to occur whether reading aloud or silently. Physiological behaviors regarding prosody have been noted by researchers studying eye movements in reading that has shown prosodic information is accessed and then applied to the text on an implicit basis (Ashby, 2006; Fodor, 1998, 2002; Kitagawa & Fodor, 2014). It has also been found that readers who encounter ambiguous syntactic structures favor syntactic replacements that allow for suitable, alternating rhythm of strong and weak syllables (Kentner, 2012, 2015; Kentner & Vasihth, 2016; McCurdy, Kentner, & Vasishth, 2013).

Using the Multi-Dimensional Fluency Scale rubric (Zutell & Rasinski, 1991) to analyze oral reading prosody, the present study investigates the development of prosodic reading in 1st-, 2nd-, and 3rd-grade students by examining the following research questions:

- What is the developmental trajectory of prosodic reading across first, second, and third grade students?
- If present, what are the differences in the rate of emergence of prosodic reading indicators?
- To what extent does prosody predict unique variance in comprehension after accounting for automaticity and accuracy?

For clarity, we use the term *accumaticity* to refer to the measurement metric reflecting automaticity and word identification accuracy, or what is often called words-correct-per-minute (WCPM). We will use the term *fluency* to refer simultaneously to the three indicators of fluent reading, that is, automaticity, word identification accuracy, *and prosody* (Samuels, 2007).

3. Method

3.1 Participants

This study was conducted in a suburban county outside a large city in the south-central U.S. Children participating in this study come from households in a county where 91.7 % of residents have graduated from high school and 40 % of adults possess a Bachelor's degree. The median household income of the district is $83,000 making it close to twice that of the state. The percentage of students receiving free- or reduced-price lunch across the study schools varied from a low of 2.8 % to a high of 52.3 % with the mean being 24.0 % and the median 21.1 %. Students attending the study schools were of primarily European-American ethnicity (91.6 %), while some were African-American (4.2 %), and the remainder came from Hispanic, Latino, and Asian populations (4.2 %). Exactly 49 and 51 percent of students enrolled in the study are female and male respectively with approximately 6.4 % diagnosed with mild to moderate learning disabilities. No students receiving

English language support services were enrolled in the study. First-grade students averaged six years three months of age at the time of the study, 2nd-grade students averaged seven years three months of age, while the mean age for 3rd-grade students was eight years two months.

3.2 Participant selection

To select participants from 10 schools involved in the study, 22 teachers were randomly selected from the 92 who instructed 1st-, 2nd-, and 3rd-grade students. All students attending the 1st-, 2nd- and 3rd-grades in these 22 classrooms were asked to return parental informed consent forms with 89 % doing so (511 students). From this pool, 100 1st-, 100 2nd-, and 100 3rd-grade students were randomly selected. At the conclusion of the study data had been fully collected for 85 1st-, 75 2nd-, and 90 3rd-grade students (n = 250).

3.3 Assessments

3.3.1 The Gray Oral Reading Test-5

The Gray Oral Reading Test-5 (GORT; Wiederholt & Bryant, 2012) is a normative, standardized assessment of oral reading proficiency obtained through a series of increasingly difficult narrative reading passages ranging between approximately 100 to 250 words. The test administrator records the number of reading word recognition miscues (mispronunciations, omitted and inserted words) and the time required to read each passage. Students read passages until a basal and ceiling level are obtained, at which point the assessment is discontinued. Separate scores are reported by the GORT-5 for word reading accuracy and automaticity. After reading each passage aloud, the student then answers a set of five factual and inferential comprehension questions. The student silently reads each question and chooses from among four multiple-choice answers. Coefficient alpha reliabilities reported by the test authors are .92 for both accuracy and automaticity, and .94 for comprehension. Test-retest coefficients equal to .90 (automaticity), .85 (accuracy), and .82 (comprehension).

3.3.2 The Multi-Dimensional Fluency Scale.

The Multi-Dimensional Fluency Scale (MDFS; Zutell & Rasinski, 1991) is a rubric for assessing fluent oral reading. The MDFS uses four indicators to evaluate a student's reading which includes the use of expression, phrasing, smoothness (accuracy in word reading), and pacing (automaticity). The quality of each indicator is rated on a scale of 1 to 4, with 4 representing the best performance. Sub-scores are summed to obtain an overall rating ranging from 4 to 16, with a score of 12 or better indicating fluent reading. The MDFS has been found to be a valid and reliable instrument in multiple studies (Paige et al., 2014; Rasinski, 1985; Rasinski et al., 2009). Moser, Sudweeks, Morrison, and Wilcox (2014) used a generalizability study (G study) of two raters where the MDFS was found to have reliability coefficients ranging from .94 to .97 for narrative text and .92 to .98 for informational text, suggesting high consistency between raters.

Reading passages from which prosody scores were evaluated using the MDFS were obtained from the GORT-5 (Wiederholt & Bryant, 2012). Passages were selected whose Lexile scores were within the CCSS (2010) Lexile grade-level bands. For the fall and spring assessments, 1st-grade students read Form A story 3 (250L), and for the winter they read story 2 (270L) from Form B (note the CCSS makes no Lexile recommendation for the 1st-grade level). 2nd-grade students read Form A, story 2 (410L), in the fall and spring, and story 4, Form B (450L) in the winter. 3rd-grade students read Form A story 4 (520L) in the fall and spring, and story 4 (530L) from Form B in the winter. Students were instructed to use their best reading voice and then asked to read aloud the full passage while being recorded on a laptop computer. To score the 750 readings gathered for the present study (each of the 250 students were assessed at three points across the school year), four graduate students in literacy were trained over two days in the use of the MDFS. Raters were led in training by one of the study authors, an expert in prosody scoring using the MDFS. A total of 50 readings were analyzed using a scaffolded training approach where raters learned to evaluate each of the four dimensions of the MDFS rubric. By the end of the second day all four raters were consistently agreeing (90 % or greater) with the expert rater. Remaining readings were then analyzed by each of the four raters and the study author with the final metric being the average of the five raters. To fully understand the variability in student scores and the factors accounting for it in the present study, a two-facet, fully crossed, generalizability study (G study) that random sampled 177 students across two of the time periods using five raters (four raters plus one of the authors of this study) was conducted (Smith & Paige, 2017). The resulting relative error variance for student by rater by time (error) was $\sigma_{\delta}^2 = 3.15$ while the generalizability coefficient was $E\,\rho^2 = .80$. The amount of variance in student scores as analyzed using the MDFS attributable to rater effects was 0. The main effect of time was also 0. Individual differences between student scores explained 63 % of the total variability. This is analogous to the true score variance among the objects of measurement in classical test theory and not a source of error (Shavelson, Webb & Noren, 1991). The interaction of student and time accounts for 29 % of the total variability indicating scores, when averaged by student and time, differ significantly. The total percent of variability in scores explained by student differences and the interaction of time and students is 93 %. The amount of unexplained variance in the scores is approximately 7 %.

3.3.3 Assessment administration.

To gather the assessment measures, a district team composed of 15 literacy coaches was assembled to administer the fluency assessment at three points (fall, winter, and spring) across the school year. Assessment administrators were trained in two 2-hour sessions. In the first session an overview of the project was given, followed by an explanation of the fluency assessments. Trainees were then put into pairs to engage in peer-to-peer administration practice. In the second session all coaches again practiced administering the GORT assessment in teams of two. Each coach administered the assessment to two different peer coaches who acted as an elementary school reader.

4. Results

4.1 Interrater reliability

The GORT-5 assessment was administered three times to approximately 250 students with all assessments graded by the researchers. To establish interrater reliability of the scoring process, a random sample of 150 assessments (20.0 %) were independently graded by two of the researchers on the measures of automaticity, word identification accuracy, and comprehension. Scores for each of the two raters were then analyzed for reliability. The resulting Cohen's kappa statistic was equal to .93 suggesting high interrater reliability commensurate with those reported above by the test authors.

4.2 Descriptive findings

Table 1 shows the means and standard deviations for the measured variables by grade level while bivariate correlations are shown in Table 2. Inspection of the means reveals increasing reading proficiency across all grades and measurement periods with the exception of the fall of second-grade where several of the measures decline from the previous spring. We compared the raw score means to the percentile rank from the GORT-5 Examiner's Manual and found that attainment for this group of students ranged between the 50th and 75th percentile depending on grade-level, indicating average to better than average reading achievement. Table 1 also reveals increases in the MDFS prosody scores across all measurement periods through the spring of 2nd-grade, by which time students had reached a score of 12, indicating fluency proficiency. After the spring of 2nd-grade, no further increases are noted until a small increase equal to .12 is found in the spring of 3rd-grade. This suggests a developmental trajectory that has reached asymptote. The table of bivariate correlations reveals statistically significant (p = .01) and general-

Table 1: Means and Standard Deviations for the Measured Variables by Grade and Assessment Period

	1st Grade (n = 85)			2nd Grade (n = 75)			3rd Grade (n = 90)		
	Fall	Winter	Spring	Fall	Winter	Spring	Fall	Winter	Spring
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
GORT Word ID Accuracy	10.81 (9.51)	16.25 (10.64)	22.77 (8.79)	24.25 (8.39)	28.89 (9.49)	30.79 (7.27)	26.81 (9.20)	32.24 (10.08)	33.60 (9.43)
GORT Automaticity	8.49 (7.03)	12.17 (8.40)	18.54 (8.34)	22.25 (9.50)	25.30 (8.77)	28.96 (7.95)	25.93 (8.65)	29.63 (9.18)	31.71 (9.58)
GORT Comprehension	10.47 (8.15)	15.13 (8.50)	19.16 (7.89)	19.49 (6.98)	24.84 (7.94)	26.49 (7.26)	22.08 (7.69)	28.58 (8.59)	26.54 (8.53)
MDFS Prosody	6.84 (4.19)	8.14 (3.62)	9.20 (3.38)	10.53 (3.35)	11.00 (2.94)	12.40 (2.80)	12.18 (3.20)	12.12 (2.97)	12.52 (2.49)
MDFS Smoothness	1.85 (1.16)	2.15 (1.11)	2.57 (1.03)	2.88 (.91)	2.81 (.82)	3.13 (.76)	3.06 (.81)	3.12 (.80)	3.17 (.66)
MDFS Pacing	1.81 (1.18)	2.24 (1.13)	2.42 (1.00)	2.87 (.89)	3.07 (.89)	3.35 (.81)	3.22 (.80)	3.33 (.83)	3.38 (.76)
MDFS Expression	1.62 (1.10)	1.74 (.94)	2.15 (.99)	2.41 (.93)	2.55 (.84)	2.97 (.82)	2.98 (.89)	2.91 (8.7)	3.00 (.76)
MDFS Phrasing	1.51 (1.08)	1.85 (.96)	1.94 (.97)	2.37 (.94)	2.57 (.77)	2.94 (.87)	3.00 (.86)	2.78 (.86)	3.01 (.79)
MDFS Smooth Pacing	1.68 (1.07)	2.19 (1.10)	2.50 (.98)	2.63 (.86)	2.94 (.81)	3.24 (.73)	3.03 (.78)	3.23 (.77)	3.27 (.64)
MDFS Expressive Phrasing	1.56 (1.06)	1.79 (.93)	2.05 (.93)	2.39 (.89)	2.56 (.77)	2.96 (.79)	2.99 (.84)	2.84 (.82)	3.01 (.71)

Note. GORT Word ID Accuracy = word identification accuracy; MDFS = Multi-Dimensional Fluency Scale

Table 2: Bivariate Correlations of the Measured Variables by Assessment Period

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. T2Accuracy	1																							
2. T1Automaticity	.964	1																						
3. T1Comprehension	.774	.784	1																					
4. T1Smoothness	.775	.811	.654	1																				
5. T1Pacing	.797	.844	.701	.917	1																			
6. T1Expression	.747	.787	.682	.795	.839	1																		
7. T1Phrasing	.749	.795	.668	.825	.857	.900	1																	
8. T1Prosody	.801	.847	.710	.929	.952	.928	.941	1																
9. T2Accuracy	.887	.877	.703	.714	.744	.706	.712	.753	1															
10. T2Automaticity	.867	.920	.726	.758	.798	.759	.782	.810	.899	1														
11. T2Comprehension	.703	.719	.732	.595	.615	.587	.615	.633	.762	.766	1													
12. T2Smoothness	.712	.748	.624	.765	.760	.676	.699	.763	.720	.775	.600	1												
13. T2Pacing	.756	.798	.664	.776	.796	.737	.724	.792	.743	.808	.645	.877	1											
14. T2Expression	.689	.734	.634	.707	.722	.778	.730	.771	.666	.734	.586	.738	.808	1										
15. T2Phrasing	.726	.757	.657	.748	.740	.760	.740	.784	.691	.741	.608	.753	.808	.884	1									
16. T2Prosody	.779	.820	.697	.808	.814	.793	.779	.838	.762	.826	.663	.909	,945	.921	.927	1								
17. T3Accuracy	.794	.794	.607	.660	.674	.619	.637	.681	.878	.835	.693	.685	.696	.594	.604	.697	1							
18. T3Automaticity	.823	.856	.660	.716	.737	.668	.691	.739	.873	.893	.719	.743	.768	.683	.694	.780	.941	1						
19. T3Comprehension	.572	.577	.647	.492	.483	.436	.424	.482	.571	.581	.666	.557	.562	.471	.537	.580	.639	.651	1					
20. T3Smoothness	.648	.673	.546	.668	.678	.597	.602	.671	.660	.687	.546	.726	.749	.644	.661	.750	.667	.731	.487	1				
21. T3Pacing	.698	.739	.606	.707	.723	.617	.645	.703	.715	.758	.642	.744	.823	.694	.681	.798	.659	.743	.511	.797	1			
22. T3Expression	.570	.606	.553	.562	.587	.593	.560	.605	.562	.618	.532	.551	.642	.681	.630	.675	.498	.574	.449	.567	.716	1		
23. T3Phrasing	.666	.698	.603	.641	.661	.634	.645	.679	.655	.698	.596	.616	.702	.725	.726	.746	.564	.650	.479	.662	.776	.796	1	
24. T3Prosody	.727	.763	.647	.725	.741	.687	.691	.747	.733	.780	.650	.740	.818	.772	.758	.834	.674	.763	.545	.844	.925	.866	.911	1

Note. All correlations significant at $p < .01$. T1 = fall assessment period; T2 = winter assessment period; T3 = spring assessment period; Accuracy = GORT word identification accuracy; Automaticity = GORT automaticity; Comprehension = GORT comprehension; Smoothness = MDFS smoothness; Pacing = MDFS pacing; Expression = MDFS expression; Phrasing = MDFS phrasing.

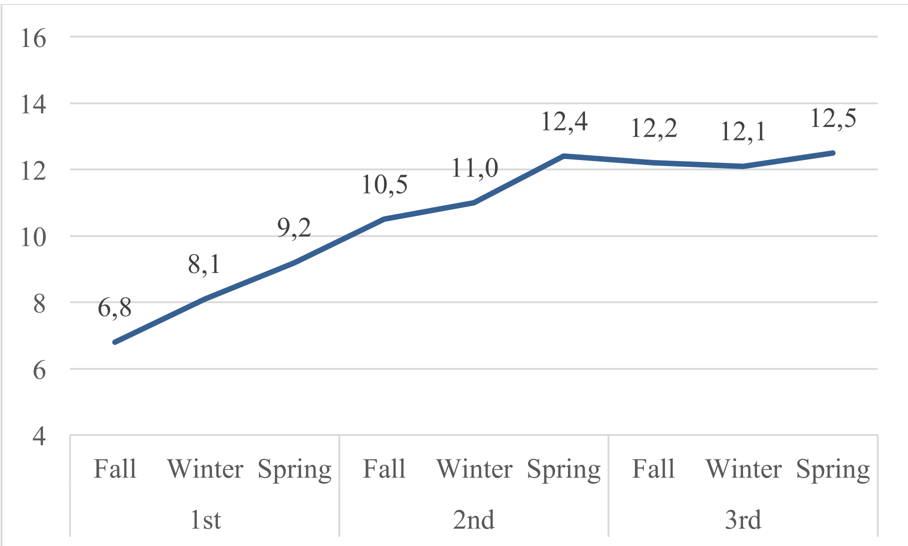
ly large relationships among the majority of variables, particularly those among the MDFS and the GORT. While still large, the correlations are smallest among comprehension and indicators of fluent reading.

4.3 Research question one

The first research question explores the developmental trajectory of prosodic reading across first-, second-, and third-grade students. To determine where statistically significant change occurs across grades and assessment periods, we used a mixed design consisting of one, 3-level within factor (time of year: fall, winter, spring) and one, 3-level between factor (grade: 1st, 2nd, and 3rd). Analysis was then conducted using repeated measures, univariate analysis of variance (ANOVA). This resulted in an analysis of prosody by time and grade, and for the simple effects of grade by time. A Bonferroni adjustment was applied to all analyses to account for potentially inflated Type I errors.

We began by testing for the assumption of sphericity with Mauchly's test indicating a violation $\chi^2(2) = 29.00$, $p < .001$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates ($\varepsilon = .90$). The test results shown in Table 3 for the main effect of time on prosody was significant, $F\ (1.8, 444.57) = 51.84$, $p < .001$, $\eta^2 = .17$, suggesting differences across the three measurement periods (Figure 1).

Figure 1: Prosody Means by Time by Grade



The test for between-grade effects (Table 3) was statistically significant, F (2, 247) = 48.26, $p < .001$, $\eta^2 = .28$ with pairwise comparisons revealing statistically significant differences between 1st- and 2nd-grade ($p < .001$), but not between 2nd- and 3rd-grade ($p = .113$). Pairwise comparisons for the effect of time on prosody show statistically significant differences between fall and winter ($p < .001$), and winter and spring ($p < .001$).

Table 3: Analysis of Variance Results for Main Effects and Interaction Effects of Time by Grade

Variable	*df*	MS	*F*	η^2
Main effect of time	(1.8, 444.57)	163.97	51.84***	.17
Main effect of grade	(2, 247)	419.24	48.26***	.28
Time x grade	(3.6, 444.57)	29.55	9.34***	.07

Note. ***p < .001.

In addition to main effects of time and grade on prosody being significant, results (Table 3) revealed that the interaction effect of time and grade was also significant, F (3.6, 444.57) = 9.34, $p < .001$, $\eta^2 = .07$. An analysis of the simple effects using pairwise comparisons with a Bonferroni adjustment revealed statistically significant differences in prosody for first-grade between fall and winter, ($p < .001$), and winter and spring ($p = .012$), and for second-grade between fall and winter ($p < .001$), but not between winter and spring ($p = .078$). For 3rd-grade a statistically significant drop in prosody ($p < .001$) was found between fall and winter, with no change between winter and spring.

4.4 Research question two

The second research question asks if the four indicators of prosodic reading (pacing, smoothness, expression, and phrasing) develop equally, or if variance exists in their emergence across grades (Figure 2).

Inspection of Table 1 shows the means for pacing and smoothness are generally larger than those for expression and phrasing, suggesting developmental differences. The bivariate correlations in Table 2 reveals the relationships between smoothness/pacing and expression/phrasing are larger than for other combinations of these four variables. Based on these findings, we hypothesized that smoothness and pacing may develop more rapidly than do expression and phrasing. To test this hypothesis we created individual composites for smoothness, pacing, expression, and phrasing. We then summed the scores for each variable across the fall, winter, and spring measures and divided by three to obtain a mean score. We then conducted paired-sample t-tests among pacing, smoothness, expression and phrasing to determine the extent to which they differed with results shown in Table 4.

Figure 2: Means for Expression, Phrasing, Smoothness, and Pacing by Grade and Assessment Period

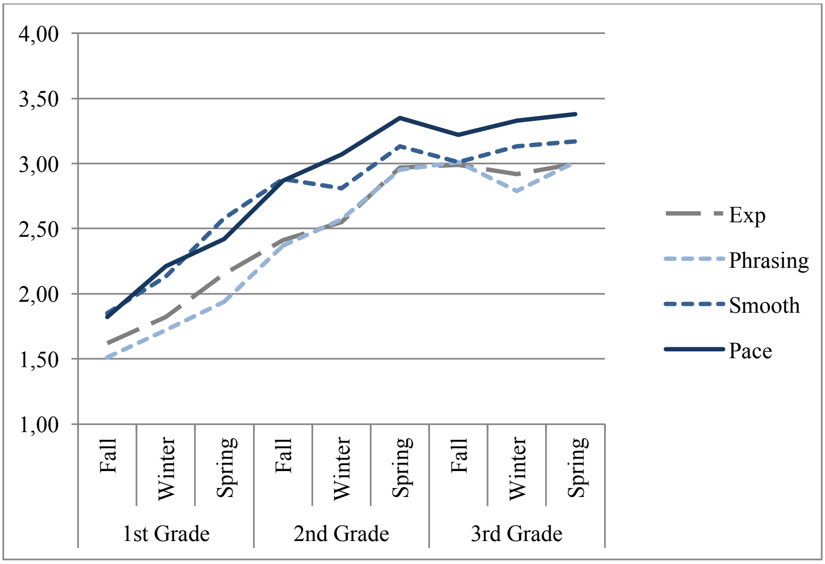


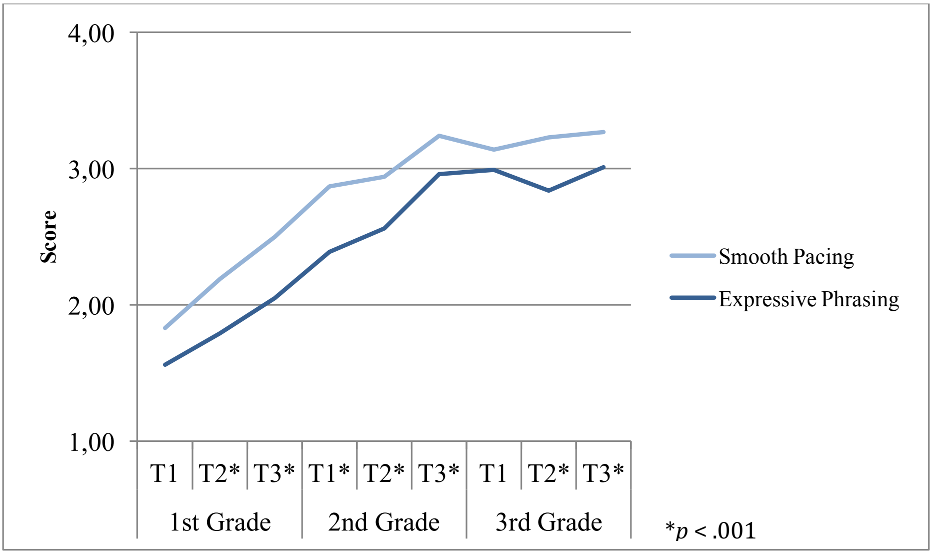
Table 4: Results for t-test Differences for Smoothness, Pacing, Expression, and Phrasing

	Smoothness *df* = 249	Pacing *df* = 249	Expression *df* = 249
Pacing	4.78*** d = .11		
Expression	7.23*** d = .28	11.30*** d = .37	
Phrasing	10.25*** d = .35	14.62*** d = .44	2.91*** d = .08

Note. ****p* < .001

After applying a Bonferroni adjustment, tests for all six comparison pairs were found to be statistically significant; however, what is more revealing are the magnitudes in the differences as measured by Cohen's *d*. Using interpretations from Hopkins (2006), the differences in magnitude for pacing and smoothness (*d* = .11) were negligible when compared to those for pacing and expression (*d* = .37, moderate), pacing and phrasing (*d* = .44, moderate), and smoothness and expression (*d* = .28, small). For the variables phrasing and expression the difference in magnitude was also negligible (*d* = .08), while differences between phrasing and smoothness (*d* = .35), phrasing and pacing (*d* = .44, moderate), expression and smoothness (*d* = .28, small) and expression and pacing (*d* = .37, moderate) were larger. With these results providing evidence that smoothness-pacing, and expression-phrasing emerge together, we proceeded to form two variables we call a) *smooth*

Figure 3: Means for Smooth Pacing and Expressive Phrasing by Grade and Assessment Period



pacing, and b) *expressive phrasing*. These two variables were formed by aggregating across grade and time to create unitary measures (see Figure 3).

To determine if smooth pacing and expressive phrasing emerged differently in readers, we conducted one paired-sample t-test of the two means. Results in Table 5 show a statistically significant difference that bordered on a small to moderate effect, t (249) = 11.42, p < .001, d = .32 revealing differences in the emergence of the two variables. We also conducted paired-sample t-tests with a Bonferroni adjustment by grade (Table 5) which resulted in statistically significant differences with small to moderate effects between smooth pacing and expressive phrasing at the 1st-, t (84) = 6.87, p < .001, d = .36, 2nd-, t (74) = 7.01, p < .001, d = .41, and 3rd-grade, t (89) = 6.01, p < .001, d = .34, levels.

4.5 Research question three

To determine the contributions of automaticity, word identification accuracy, smooth pacing, and expressive phrasing to reading comprehension, we first created composite variables for each variable by summing scores across the fall, winter, and spring measures, and then divided by three. We then regressed the four variables onto comprehension: automaticity, word identification accuracy, smooth pacing and expressive phrasing. Initial findings showed that smooth pacing and automaticity were not significant predictors so both were eliminated from the model.

Table 5: Means (sd) and t-test Results for Differences Between Smooth Pacing and Expressive Phrasing by Grade.

		Grade		
Variable	Total	1	2	3
Smooth Pacing	2.75 (.91)	2.12 (.95)	2.94 (.73)	3.18 (.66)
Expressive Phrasing	2.46 (.90)	1.80 (.85)	2.64 (.73)	2.95 (.69)
t	11.42***	6.87***	7.01***	6.01***
(df)	(249)	(84)	(74)	(89)
d	.32	.36	.41	.34

Note. ***$p < .001$. d = Cohen's d.

Table 6: Hierarchical Regression Analysis for Fluency Measures Predicting Comprehension

	B	SE β	β	R^2	ΔR^2	t
Constant	3.68	.94				
Word Identification Accuracy	.461	.049	.582	.624		9.33***
Expressive Phrasing	2.46	.582	.264	.649	.025	4.23***

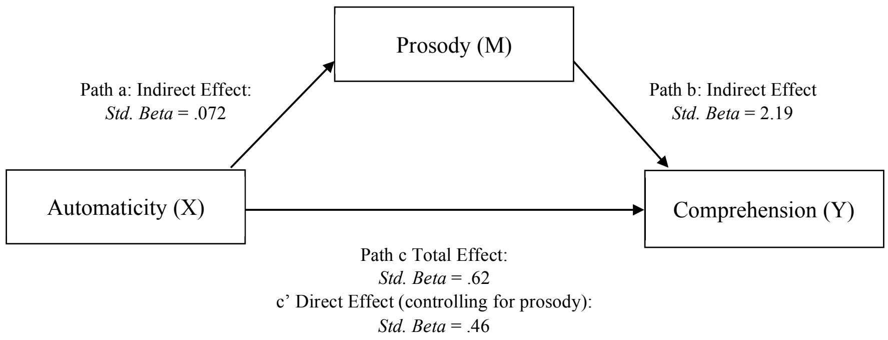
Note. ***$p < .001$.

Final results in Table 6 show that word identification accuracy and expressive phrasing explained 64.9 % of the variance in comprehension, F (2,240) = 223.06, $p < .001$, $d = 1.36$. Beta coefficients were equal to .582 ($t = 9.33$, $p < .001$) for accuracy and .264 ($t = 4.23$, $p < .001$) for expressive phrasing. Word identification accuracy accounted for 62.4 % of the variance while expressive phrasing contributed an additional 2.5 % to comprehension.

4.6 Mediation analysis

Prosody has been found to mediate the relationship between accumaticity and comprehension in secondary students (Paige et al., 2014) and the result that accumaticity was not a significant predictor of comprehension leaves open the possibility that this same phenomenon is occurring in the present study. To test this possibility we used a three-step mediation analysis (Baron & Kenny, 1986) with bootstrapping which has been found to increase the stability of parameter estimates (Preacher & Hayes, 2004; Preacher & Kelley, 2011). To show a mediation effect, Baron and Kenny state that variable X must be correlated with both Y and M, and M must be correlated with Y. Additionally, M must be correlated with Y while controlling for X. Finally, the direct effect of X on Y must be eliminated or reduced

Figure 4: Pathway of Mediation Analysis



when Y is regressed on M and X. The path model defining the relationships among the tested variables is shown in Figure 4.

We began by regressing expressive-phrasing onto automaticity (path a), then comprehension onto expressive-phrasing (path b), followed by comprehension onto automaticity (path c).

Table 7: Regression Results for the Mediation of the Effect of Automaticity on Comprehension by Prosody

Model/(path)	Estimate	SE	95 % CI (lower)	95 % CI (upper)
AUTO – PROS (a) $R^2_{M.X}$	.072*** .708	.003	.066	.078
PROS – COMP (b)	2.19***	.685	.842	3.542
AUTO – COMP (c') $R^2_{Y.MX}$	.462*** .624	.032	.347	.577
AUTO – COMP (c) $R^2_{Y.X}$	.619*** .608	.032	.556	.683

Note. Estimates for model/path effects are standardized coefficients (betas). $R^2_{M.X}$ is the proportion of variance in prosody (M) explained by automaticity (X). $R^2_{Y.MX}$ is the proportion of variance in comprehension (Y) explained by prosody (M) and automaticity (X). $R^2_{Y.X}$ is the proportion of variance in comprehension (Y) explained by automaticity (X). CI = confidence interval.
***$p < .001$.

Results in Table 7 show that the regression model meets the Baron and Kenny criteria to state that prosody asserts a partial, mediating effect on the relationship between automaticity and comprehension. Table 7 shows the beta coefficients associated with each of these conditions were all significant at $p < .001$. The effect of automaticity on comprehension is reduced from .619 to .462 when expressive-phrasing is added to the equation which makes the effect of prosody measured by Cohen's kappa of $\kappa^2 = .157$ (Preacher & Kelley, 2011), a moderate-sized effect (Hopkins, 2006). The reader should be aware that Preacher and Kelley recommend caution in applying a qualitative descriptor to effect sizes.

To account for possible weaknesses in the Baron and Kenny model, we conducted a resampling approach to increase the stability of the parameter estimates (Preacher & Hayes, 2004; Preacher & Kelley, 2011). Table 7 shows the 95 % confidence intervals calculated through SPSS's bias corrected bootstrap for total, direct, and indirect effects of the mediation model. As the confidence intervals do not include 0, it can be concluded that all effects are significant considering a null hypothesis of $b = 0$. To test the indirect effect (ab) for significance we used the normal theory test which is available as an add-on developed by Hayes (2013). Results of both the traditional approach (Baron & Kenny, 1986) and the more robust bootstrap estimation (Hayes, 2013) indicate that prosody exerts a significant mediating effect on the relationship between automaticity and comprehension.

5. Discussion

The present study measured word reading accuracy, automaticity, prosody, and comprehension in a randomly selected group of 250 1st-, 2nd-, and 3rd-grade students from a suburban, upper-middle class school district. Our goal was to first, identify the developmental trajectory of prosodic reading behaviors across grades one to three. We next sought to determine if differences existed in the emergence of prosodic reading indicators, and finally we analyzed the extent to which the three indicators of fluent reading predicted reading comprehension. We found that by the end of 2nd-grade, students had achieved a criterion of 12 (maximum score = 16) on the MDFS, a score suggesting acceptable prosodic reading, with no subsequent growth throughout the 3rd-grade. This result suggests that students had reached asymptote in prosodic reading by the end of second grade. It is important to remember that the MDFS has a range of 4 to 16, so our results do not represent ceiling effects due to constraints of the measurement instrument. Simple effects analysis revealed that assessment time (fall, winter, spring), grade level, and the interaction of the two, explained growth in reading prosody.

We were also interested in determining if differences existed in the rate of emergence of the four MDFS indicators associated with prosodic reading. After collapsing the four indicators into two variables we named smooth-pacing (akin to accumaticity) and expressive-phrasing, we found that students progressed more quickly in the development of smooth-pacing than they did in expressive-phrasing. Finally, we wanted to know the extent to which comprehension variance was explained by the measures of word identification accuracy, automaticity, smooth-pacing, and expressive-phrasing. Results revealed that accumaticity and smooth-pacing were not significant predictors of comprehension, while word identification accuracy and expressive-phrasing (prosody) accounted for 64.9 % of comprehension. Our results further emphasize the importance of reading prosody by a) describing the emergence of prosodic reading development in early elementary students; b) providing empirical evidence that young readers progress more quickly in their development of reading automaticity and accuracy than in the application of expression and phrasing; and c) showing that prosody in the form of expressive-phrasing contributes unique variance to reading comprehension in early elementary students.

These results provide a comparison to those of Schwanenflugel et al. (2004). In their study of 1st- through 3rd-grade students drawn primarily from lower-SES households, the authors measured numerous reading features contributing to automaticity and fluent reading development. In contrast to ours, their results did not find that connected text reading predicted comprehension, rather, they support a *simple reading fluency* model where text reading fluency functions as an additional indictor of fluent reading. On the other hand, our results strongly support connected text reading as a predictor of reading comprehension and specifically illuminates the role of prosody in sentence-parsing skills (Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003; Schreiber, 1980, 1987). We agree with Schwanenflugel et al. (2004) who suggest the developmental level of the reader must be considered in evaluating the contribution of prosodic, connected-text reading to comprehension. Studies reviewed earlier by Veenendaal et al. (2014, 2015) exploring prosodic reading in Dutch 4th-grade students with proficient levels of fluency also found results similar to ours that support the notion that reading proficiency is important to the emergence of prosodic reading. Our sample of students with fluent reading skills, as well as those from the Veenendaal et al. studies, suggest that decoding development acts as a foundation for prosodic reading in students possessing proficient decoding skills.

Of importance to our findings is an understanding of how to interpret the unique contribution of prosody to reading comprehension. We found that automaticity, or reading pace, was not a significant predictor of reading comprehension, which puts our results in alignment with those of Veenendaal et al. (2014) and Valencia et al. (2010). Our mediation analysis found that expressive reading accounted for a significant amount of the variance between automaticity and comprehension, further supporting the idea that decoding proficiency is important to support prosodic reading. Our finding that rate or automaticity was not a significant predictor of reading comprehension also provides support for the tandem theory (Paige et al., 2014) that hypothesizes how the indicators of fluent reading interact to support comprehension. Our results show the average student in this study read with acceptable prosody and average or better rate, word identification accuracy, and comprehension. The fact that rate was under control by the reader and did not predict comprehension suggests students were able to adjust rate in relation to comprehension while using prosodic reading skill in the form of expressive phrasing to assist with comprehension. Our results also provide some support for the implicit prosody hypothesis (IPH) that suggests when confronted with ambiguous text, the reader may project several different prosodic contours onto the text, in a trial and error fashion, choosing the one that seems most natural (Fodor, 2002). In so doing, this process may influence or be helpful to the reader in resolving ambiguous syntax, thus aiding comprehension. Critical to the IPH hypothesis is that the reader possess sufficiently developed prosodic reading skills to engage in the process. Our results show that a) students in the study had well developed prosodic reading ability and b) that prosody contributes to explaining unique variance in reading comprehension beyond that of other fluency indicators. Our assessment of prosody considered the extent to which the reader applied it across the entire reading passage, not just too a specific phrase or sentence, while the IPH addresses the use of prosody in specific instances of textual ambiguity. As such, prosody may contribute to textual understanding in potentially two ways. The first as suggested by other researchers reviewed earlier, is in the creation of an overall or macro-level cognitive architecture that functions as a foundation for understanding. The second use of prosody is on an occasional basis where it is employed as a problem-solving tool used to resolve ambiguous meaning at the phrase level.

If prosody is important on both a macro- and micro-level, then how should we interpret the 2.5 % of explained variance in comprehension that by Cohen's (1988) rule-of-thumb interpretation is trivial in size? To avoid misleading conclusions Cohen himself, as well as multiple researchers since, suggest that applying a rule-of-thumb interpretation of effect sizes may result in the dismissal of important results (Hill, Bloom, Black, & Lipsey, 2008; Kane, 2004; McCartney & Rosenthal, 2000; Prentice & Miller, 1992). Hill et al. (2008) suggest that the research base regarding the variable of interest, the research context, and the participant subgroup of interest are some of the factors to be considered in effect size interpretation. The research base investigating the effect of prosody on comprehension is still emerging as relatively few studies have been conducted with some finding statistically significant effects for prosody while others have found none. Secondly, the tools for measuring prosody vary according to the academic orientation of the researcher, while analysis of results also differ from correlation and hierarchical regression analyses, to other designs. Moreover, the limited range (4-16) of scores possible with the prosody metric used in the present study restricts the possibility of more robust relationships between prosody and comprehension. Third, the studies that have found statistically significant effects of prosody on comprehension have reported effect sizes in the single-digit range. Fourth, in addition to the present study, one other analyzing secondary students (Paige et al., 2014) has found that prosody mediates a significant amount of variance between automaticity and comprehension, a finding that further specifies the effect of prosody. While few individuals wish to listen to anyone read aloud in a robotic voice devoid of prosody, this fundamental level of prosodic reading may prepare the stage for its use as a problem-solving tool. We also suggest that readers with generally acceptable decoding proficiency who effectively monitor their comprehension have sufficiently developed reading skills to invoke the use of prosody as a problem-solving strategy. As such, the importance of its small effect size is suggestive not so much of its relative importance to overall comprehension, but rather, to the critical importance of prosody as a problem-solving strategy for use with ambiguous text when nothing else will suffice. In such cases the seemingly small effect size of prosody may defy its critical importance to the reader.

The finding of prosody's contribution to reading comprehension suggests that prosody should be part of an effective reading instruction curriculum in the primary grades. Simple instructional activities such as teachers' modeling and explanation of expressive reading to students and setting up instructional scenarios in which students engage in repeated readings of texts in order to improve their expressive oral interpretation of the texts may lead to increases in overall reading achievement. Indeed, classroom-based studies have shown that such prosody-focused instruction can have positive effects on achievement in students through grade 4 (e.g., Griffith & Rasinski, 2004; Martinez, Roser, & Strecker, 1999; Young & Rasinski, 2009).

5.1 Study limitations

The student population in this study was randomly drawn from primarily upper-middle class households; as such, our results do not consider those from generally lower-SES backgrounds. Because of this we caution the reader in the interpretation of our results. Our study is also limited to a single geographic location. The inclusion of students from multiple locations of the country may yield different results. The reader should be aware that the MDFS is just one way of measuring prosody and that other methodologies may, and in fact have, yielded different results. Other methodologies include prosody analysis using speech software that can analyze and measure speech data in ways not possible using reading rubrics. While the rubric used in this study (MDFS) has been shown to be highly reliable when raters are properly trained, use of speech analysis software may have yielded different results.

5.2 Future research

The first question concerns the maintenance of prosodic reading as students' progress through grades. In other words, as text complexity increases, are prosodic reading behaviors developed in elementary school maintained in upper grades and how much reading is required to maintain fluent reading? Another question asks what outcomes might be obtained for a group of readers who struggle with reading acquisition, and what is the trajectory of their prosody development? Does the maintenance of prosodic reading require different amounts of time-spent-reading depending upon the extent to which the reader exhibits fluent reading? To what extent does the maintenance of reading prosody depend on the amount of continued practice with increasingly complex text? Finally, what factors determine when or if, readers from lower-SES backgrounds achieve prosodic reading?

References

ACT. (2013). *The condition of college and career readiness 2014*. Retrieved from http://www.act.org/research/policymakers/cccr13/pdf/CCCR13-NationalReadinessRpt.pdf

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Ashby, K. (2006). Prosody in skilled silent reading: Evidence from eye movements. *Journal of Research in Reading, 29*(3), 318–333.

Ashby, J., & Clifton, C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition, 96*(3), B89–B100. doi: 10.1016/j.cognition.2004.12.006

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321–324.

Bader, M. (1998). Prosodic influences of reading syntactically ambiguous sentences. In J. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing* (pp. 1–46). Dordrecht, Netherlands: Kluwer.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182.

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes, 11*(1-2), 17–67.

Benjamin, R. G., & Schwanenflugel, P. J. (2010). Text complexity and oral reading prosody in young children. *Reading Research Quarterly, 45*(4), 388–404. doi: dx.doi.org/10.1598/RRQ.45.4.2

Breznitz, Z. (2006). *Fluency in reading: Synchronization of processes*. Mahwah, NJ: Erlbaum.

Carlson, K. (2009). How prosody influences sentence comprehension. *Language and Linguistics Compass, 3*(5), 1188–1200.

Chafe, W. (1988). Punctuation and the prosody of written language. *Written Communication, 5*(4), 395–426. doi: 10.1177/0741088388005004001

Christopher, A., Mehler, J., & Sebastian-Galles, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy, 2*(3), 385–394. doi: 10.1111/j.1749-818x.2008.00103.x

Church, K. (1987). Phonological parsing and lexical retrieval. *Cognition*, *25*(1), 53–69.

Cole, R. A., & Jakimik, J. (1980). How are syllables used to recognize words? *The Journal of the Accoustical Society of America, 67*(3), 965–970.

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

CCSS – Common Core State Standards Initiative. (2010). *Common core state standards for English language arts & literacy in history/social studies, science, and technical subjects*. Washington, DC: CCSSO & National Governors Association. Center for Reading Research Technical Report #1. Retrieved from http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Compton, D. L., Appleton, A. C., & Hosp, M. K. (2004). Exploring the relationship between text-leveling systems and reading accuracy and fluency in second-grade students who are average and poor decoders. *Learning Disabilities Research and Practice, 19*(3), 176–184.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language, 31*(2), 218–236.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language, 2*, 133–483. doi: 10.1037/0012-1649.39.4.706

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*(2), 141–201.
Daane, M. C., Campbell, J. R., Grigg, W. S., Goodman, M. J., & Oranje, A. (2005). *Fourth-Grade students reading aloud: NAEP 2002 special study of oral reading*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.
Dickinson, D. K., & Neuman, S. D. (Eds.). (2006). *Handbook of early literacy research* (Vol. 2). New York, NY: Guilford.
Dowhower, S. L. (1987). Effects of repeated reading on second-grade transitional readers' fluency and comprehension. *Reading Research Quarterly*, *22*(4), 389–406. doi: 10.2307/747699
Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory Into Practice*, *30*(3), 165–175.
Fodor, J. D. (1998). Parsing to learn. *Journal of Psycholinguistic Research, 27*(3), 339–375.
Fodor, J. D. (2002, April). *Psycholinguistics cannot escape prosody*. Paper presented at Speech Prosody 2002, Aix-en-Provence, France. Retrieved from https://www.researchgate.net/profile/Janet_Fodor/publication/2891257_PsycholinguisticsCannot_Escape_Prosody/links/09e41507c2300f3cca000000.pdf
Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences, 10*(6), 244–249.
Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*(1), 71–99. doi: 10.1037/0033-2909.123.1.71
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*(3), 239–256.
Goldman, S. R., Meyerson, P. M., & Cote, N. (2006). Poetry as a mnemonic prompt in children's stories. *Reading Psychology, 27*(4), 345–376.
Goswami, U. (2005). Synthetic phonics and learning to read: A cross-language perspective. *Educational Psychology in Practice, 21*(4), 272–282. doi:10.1080/02667360500344823
Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., & Scott, S. K. (2002). Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proceedings of the National Academy of Sciences of the United States of America, 99*(16), 10911–10916.
Griffith, L. W., & Rasinski, T. V. (2004). A focus on fluency: How one teacher incorporated fluency with her reading curriculum. *The Reading Teacher, 58*(2), 126–137.
Gross, J., Millett, A. L., Bartek, B., Bredell, K. H., & Winegard, B. (2014). Evidence for prosody in silent reading. *Reading Research Quarterly*, *49*(2), 189–208. doi: 10.1002/rrq.66
Guitérrez-Palma, N., & Palma-Reyes, A. (2008). On the use of lexical stress in reading Spanish. *Reading & Writing, 21*(6), 645–660.
Hasbrouck, J., & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, *59*(7), 636–644.
Hayes, A. F. (2013). *PROCESS for SPSS and SAS* (version 2.041). Retrieved from http://www.processmacro.org
Heilman, A. J., Blair, T. R., & Rupley, W. H. (Eds.). (2002). *Principles and practices of teaching reading*. Columbus, OH: Merrill Publishing.
Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172–177.
Himmelmann, N. P., & Ladd, D. R. (2008). Prosodic description: An introduction for field workers. *Language Documentation & Conservation, 2*(2), 244–27.

Hirotani, M., Frazier, L., & Rayner, K. (2006). Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *Journal of Memory and Language, 54*(3), 526–544.
Holliman, A., Critten, S., Lawrence, T., Harrison, E., Wood, C., & Hughes, D. (2014). Modeling the relationships between prosodic sensitivity and early literacy. *Reading Research Quarterly, 49*(4), 469–482. doi: 10.1002/rrq.82. doi: 10.1037/a0018049
Holliman, A. J., Wood, C., & Sheehy, K. (2010). Does speech rhythm sensitivity predict children's reading ability 1 year later? *Journal of Educational Psychology, 102*(2), 356–366. doi: 10.1037/a0018049
Hopkins, W. G. (2006). *A scale of magnitudes for the effect statistics: A new view of statistics.* Retrieved from http://www.sportsci.org/resource/stats/effectmag.html
Hudson, R. F., Pullen, P. C., Lane, H. B., & Torgesen, J. K. (2009). The complex nature of reading fluency: A multidimensional view. *Reading and Writing Quarterly, 25*(1), 4–32. doi:10.1080/10573560802491208
Huey, E. G. (1968). *The psychology and pedagogy of reading: With a review of the history of reading and writing and the methods, texts, and hygiene in reading.* Cambridge, MA: MIT Press.
Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95*(4), 719–729.
Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology, 29*(1), 1–23.
Jusczyk, P. W., Cutler, A., & Redanz, N. (1993). Preference for the predominate stress patterns of English words. *Child Development, 64*(3), 675–687.
Jusczyk, P. W., & Hohne, E. A. (1997). Infants' memory for spoken words. *Science, 277*(5334), 1984–1986. doi: 10.1126/science.277.5334.1984
Kane, T. J. (2004). *The impact of after-school programs: Interpreting the results of four recent evaluations* (William T. Grant Foundation Working Paper). New York, NY: William T. Grant Foundations.
Kenter, G. (2012). Linguistic rhythm guides parsing decision in written sentence comprehension. *Cognition, 123*(1), 1–20. doi: 10.1016/j.cognition.2011.11.012
Kentner, G. (2015). Stress clash hampers processing of noncanonical structures in reading. In R. Vobel & R. van de Vijver (Eds.), *Rhythm in Cognition and Grammar: A Germanic Perspective* (pp. 111–135). Berlin, Germany: De Gruyter. doi: 10.1515/9783110378092.111
Kenter, G., & Vasishth, S. (2016). Prosodic focus marking in silent reading: Effects of discourse context and rhythm. *Frontiers in Psychology, 7*(319), doi: 10.3389/fpsyg.2016.00319
Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 163–182.
Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge, United Kingdom: Cambridge University Press.
Kitagawa, Y., & Fodor, J. D. (2014). Prosodic influence on syntactic judgments. In G. Fanselow, C. Fery, R. Vogel, & M. Schlesewsky (Eds.), Gradience in grammar: Generative Perspectives (pp. 336-358). Oxford, England: Oxford University Press.
Klauda, S. L., & Guthrie, J. T. (2008). Relationships of three components of reading fluency to reading comprehension. *Journal of Educational Psychology, 100*(2), 310–321.
Koriat, A., Greenberg, S. N., & Kreiner, H. (2002). The extraction of structure during reading: Evidence from reading prosody. *Memory & Cognition, 30*(2), 270–280.
Kuhn, M. R., Schwanenflugel, P. J., & Meisinger, E. B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly, 45*(2), 230–251.

Kuhn, M. R., & Stahl, S. A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95*(1), 3–21.
LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*(2), 293–323.
Laver, G. P. M. (1997). *Principles of phonetics*. Cambridge, United Kingdom: Cambridge University Press.
Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*(4), 492–527.
Logan, G. D., & Etherton, J. L. (1994). What is learned during automatization? The role of attention in constructing an instance. *Journal of Experimental Psychology, 20*(5), 1022–1050.
Mandel, D. R., Jusczyk, P., & Nelson, D. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition, 53*(2), 155–180.
Mandel, D. R., Jusczyk, P., & Pisoni, D. (1995). Infants' recognition of sound patterns of their own names. *Psychological Science, 6*(5), 315–318.
Martinez, M., Roser, N., & Strecker, S. (1999). "I never thought I could be a star": A readers theatre ticket to reading fluency. *The Reading Teacher, 52*(4), 326–334.
McCartney, D., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. *Child Development, 71*(1), 173–180.
McCurdy, K., Kentner, G., & Vasishth, S. (2013). Implicit prosody and contextual bias in silent reading. *Journal of Eye Movement Research, 6*(2), 1–14.
MetaMetrics. (2012*). Linking the K-PREP reading with the Lexile framework: A study to link the Kentucky performance rating for educational progress reading with the Lexile framework for reading*. Retrieved from http://education.ky.gov/aa/kts/documents/k-preplexilelinkingstudy11-2012%20v2.pdf
Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, N., Bertoncini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*(2), 143–178.
Miller, J. M., & Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology, 98*(4), 839–853. doi: 10.1037/0022-0663.98.4.839
Miller, J. M., & Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly, 43*(4), 336–354. doi: 10.1598/RRQ.43.4.2
Moser, G. P., Sudweeks, R. R., Morrison, T. G., & Wilcox, B. (2014). Reliability of children's expressive reading. *Reading Psychology, 35*(1), 58–79. doi: 10.1080/02702711.2012.67541
National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. (National Institute of Health Pub. No. 00-4769). Washington, DC: National Institute of Child Health and Human Development.
Paige, D. D. (2011). Engaging struggling adolescent readers through situational interest: A model proposing the relationships among extrinsic motivation, oral reading fluency, comprehension, and academic achievement. *Reading Psychology, 32*(5), 395–425. doi: 10.1080/02702711.495633
Paige, D. D., Rasinski, T. V., Magpuri-Lavell, T., & Smith, G. S. (2014). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research, 46*(2), 123–156. doi: 10.1177/1086296X14535170
Peppé, S. J. E., Martínez-Castilla, P., Coene, M., Hesling, I., Moen, I., & Gibbon, F. (2010). Assessing prosodic skills in five European languages: Cross-Linguistic differences in typical and atypical populations. *International Journal of Speech-Language Pathology, 12*(1), 1–7. doi: 10.3109/17549500903093731

Peppé, S., & McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: The PEPS-C test and the revised version. *Clinical Linguistics & Phonetics, 17*(4–5), 345–354. doi: 1080/0269929931000079999

Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford Press.

Perfetti, C. A. (1988). Verbal efficiency in reading ability. In M. Daneman, G. E. MacKinnon, & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (pp. 109–143). New York, NY: Academic Press.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 167–2008). New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780198507932.0006

Pinnell, G. S., Pikulski, J. J., Wixon, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). *Listening to children read aloud*. Washington, DC: Office of Educational Research and Improvement, U.S. Department of Education.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*(4), 717–731. doi 10.3758/BRM.40.3.879

Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods, 16*(2), 93–115. doi: 10.1037/a0022658

Prentice, D. A., & Miller, D. T. (1992). When small effects are important. *Psychological Bulletin, 112*(1), 160–164.

Rasinski, T. V. (1985). *A study of factors involved in reader-text interactions that contribute to fluency in reading* (Unpublished doctoral dissertation). The Ohio State University, Columbus.

Rasinski, T. V., Reutzel, D. R., Chard, D., & Linan-Thompson, S. (2011). Reading fluency. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerbach (Eds.), *Handbook of Reading Research* (Vol. 4, pp. 286–319). New York, NY: Routledge.

Rasinski, T., Rikli, A., & Johnston, S. (2009). Reading fluency: More than automaticity? More than a concern for the primary grades? *Literacy Research and Instruction, 48*(4), 350–361.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *Quarterly Journal of Experimental Psychology, 53*(4), 1061–1080.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton, C. (1989). Eye movement and on-line language comprehension processes. *Language and Cognitive Processes, 4*(3-4), 21–49.

Richardson, U., Thomson, J., Scott, S., & Goswami, U. (2004). Auditory processing skills and phonological representation in dyslexic children. *Dyslexia, 10*(3), 215–233.

Samuels, S. J. (2006). Toward a model of reading fluency. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 22–46). Newark, DE: International Reading Association.

Samuels, S. J. (2007). The DIBELS tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly, 42*(4), 563–566.

Samuels, S. J., & Flor, R. F. (1997). The importance of automaticity for developing expertise in reading. *Reading & Writing Quarterly, 13*(2), 107–121. doi: 10.1080/1057356970130202

Sanderman, A. A., & Collier, R. (1997). Prosodic phrasing and comprehension. *Language and Speech, 40*(44), 391–409.

Savill, N., Lindell, A., Booth, A., West, G., & Thierry, G. (2011). Literate humans sound out words during silent reading. *NeuroReport, 22*(3), 116–120. doi: 10.1097/WNR.0b013e328242ed27

Schatschneider, C., Buck, J., Torgesen, J. K., Wagner, R. K., Hassler, L., Hecht, S., & Powell-Smith, K. (2004). *A multivariate study of factors that contribute to in-*

dividual differences in performance on the Florida Comprehensive Reading Assessment Test (Technical Report No. 5). Tallahassee, FL: Florida Center for Reading Research.

Schneider, W. (1985). Toward a model of attention and the development of automatic processing. In M. I. Posner & O. S. Marin (Eds.), *Attention and performance XI* (pp. 474–492). Hillsdale, NJ: Erlbaum.

Schneider, W., & Shiffrin, R. (1977). Controlled and automatic human information processing: 1. Detection, search, and attention. *Psychological Review, 84*(1), 1–66.

Schreiber, P. A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior, 12*(3), *177–186*.

Schreiber, P. A. (1987). Prosody and structure in children's syntactic processing. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 243–270). New York, NY: Academic Press.

Schreiber, P. A. (1991). Understanding prosody's role in in reading acquisition. *Theory into Practice, 30*(3), 158–164.

Schwanenflugel, P. J., & Benjamin, R. G. (2016). The development of reading prosody and its assessment. In J. Thomson & L. Jarmulowicz (Eds.), *Linguistic rhythm and literacy* (IASCl series' Trends in Language Acquisition Research (TiLAR), pp. 187–213). Amsterdam, Netherlands: John Benjamins.

Schwanenflugel, P. J., & Benjamin, R. G. (2017). Lexical prosody as an aspect of reading fluency. *Reading and Writing, 30*(1), 143–162. doi: 10.1007/s11145-016-9667-3

Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology, 96*(1), 119–129. doi 10.1037/0022-0663.96.1.119

Schwanenflugel, P. J., Meisinger, E. B., Wisenbaker, J. M., Kuhn, M. R., Strauss, G. P., & Morris, R. D. (2006). Becoming a fluent and automatic reader in the early elementary years. *Reading Research Quarterly, 41*(4), 406–522. doi: 10.1598/RRQ41.4.4

Schwanenflugel, P. J., Westmoreland, M. R., & Benjamin, R. G. (2015). Reading fluency skill and the prosodic marking of linguistic focus. *Reading & Writing, 28*(2), 9–30.

Shavelson, R., & Webb, N. M. (1991). *Generalizability theory: A primer*. Retrieved from https://web.stanford.edu/dept/SUSE/SEAL/Reports_Papers/methods_papers/G%20Theory%20AERA.pdf

Shriffrin, R. M., & Schneider, W. (1977). Controlled and automatic human processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(2), 127–190.

Slowiaczek, M. L., & Clifton, C. Jr. (1980). Subvocalization and reading for meaning. *Journal of Verbal Learning and Verbal Behavior, 19*(5), 573–582.

Smith, G. S., & Paige, D. D. (2017). *Applying a generalizability study of interrater reliability across multiple raters when using the NAEP and MDFS rubrics to measure oral reading fluency*. Manuscript in review.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York, NY: Guilford Press.

Stecker, S. K., Roser, N. L., & Martinez, M. G. (1998). Understanding oral reading fluency. In T. Shanahan & F. V. Rodriquez-Brown (Eds.), *47th yearbook of the National Reading Conference* (pp. 295–310). Chicago, Il: National Reading Conference.

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General, 136*(1), 64–81. doi: 10.1037/0096-3445.136.1.64

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month old infants. *Developmental Psychology, 39*(4), 706–716.

Torgesen, J. K., & Hecht, S. A. (1996). Preventing and remediating reading disabilities: Instructional variables that make a difference. In M. F. Graves, P. van den Broek, & B. M. Taylor (Eds.), *The first R: Every child's right to read* (pp. 133–158). New York, NY: Teachers College Press.

Torgesen, J. K., & Hudson, R. F. (2006). Reading fluency: Critical issues for struggling readers. In S. J. Samuels & A. Farstrup (Eds.), What research has to say about fluency instruction (pp. 130-158). Newark, DE: International Reading Association.

Torgesen, J. K., Rashotte, C. A., & Alexander, A. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, fluency, and the brain* (pp. 333–356). Timonium, MD: York Press.

Valencia, S. W., Smith, A., Reece, A. M., Li, M., Wixon, K. K., & Newman, H. (2010). Oral reading fluency assessment: Issues of content, construct, criterion, and consequent validity. *Reading Research Quarterly, 45*(3), 270–291. doi: 10.1598/RRQ.45.3.1

Van Orden, G. C., & Goldinger, S. D. (1994). The interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance, 20*(6), 1269–1291.

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2014). The role of speech prosody and text reading prosody in children's reading comprehension. *British Journal of Educational Psychology, 84*(4), 521–536. doi: 10.1111/bjep.12036

Veenendaal, N. J., Groen, M. A., & Verhoeven, L. (2015). What oral reading fluency can reveal about reading comprehension. *Journal of Research in Reading, 38*(3), 213–225. doi: 10.11111467-9817.12024

Vihman, M. M. (2014). *Phonological development: The first two years* (2nd ed.). Chichester, United Kingdom: John Wiley and Sons.

Vihman, M. M., & Velleman, S. L. (2000). The construction of first phonology. *Phonetica, 57*, 255–266.

Wade-Wooley, L. (2016). Prosodic and phonemic awareness in children's reading of long and short words. *Reading and Writing, 29*(3), 371–382. doi: 10.1007/s11145-015-9600-1

Weber, C., Hahne, A., Friedrich, M., & Friederici, A. D. (2004). Discrimination of word stress in early infant perception: Electrophysiological evidence. *Cognitive Brain Research, 18*(2), 149–161. doi:10.1016/j.cogbrainres.2003.10.001.

Wiederholt, J. L., & Bryant, B. R. (2012). *Gray oral reading test* (5th ed.). Austin, TX: Pro-Ed.

Whalley, K., & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading, 29*(3), 288–303.

Wood, C. (2006). Metrical stress sensitivity in young children and its relationship to phonological awareness and reading. *Journal of Research in Reading, 29*(3), 270–287. doi:10.1111/j.1467-9817.2006.00308.x

Young, A. R., Bowers, P. G., & MacKinnon, G. E. (1996). Effects of prosodic modeling and repeated reading on poor readers' fluency and comprehension. *Applied Psycholinguistics, 17*(1), 59–84.

Young, C., & Rasinski, T. (2009). Implementing readers theatre as an approach to classroom fluency instruction. *The Reading Teacher, 63*(1), 4–13.

Zutell, J., & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice, 30*(3), 211–21