THE DEVELOPMENT OF PROSODIC TEXT READING AS A DIMENSION OF ORAL READING FLUENCY IN EARLY ELEMENTARY SCHOOL CHILDREN

by

JUSTIN MILLER

(Under the Direction of Paula J. Schwanenflugel and A. Michele Lease)

ABSTRACT

Prosodic text reading is widely considered one of the defining characteristics of oral reading fluency. The general purpose of this study was to examine the development of reading prosody throughout the years of primary reading acquisition (grades 1 through 3). Participants were 92 first-grade students who were part of a larger study of the development of reading fluency. Suprasegmental features of oral reading were measured on three separate occasions throughout grades 1 and 2 (initial measurements were taken during the spring of the first grade school year with follow-up assessments occurring during the fall and the spring of second grade). A final outcome assessment was included during the spring of the third grade school year as well. Outcome measures consisted of formal assessments of oral reading fluency and reading comprehension. The specific research objectives consisted of the following: (1) to determine whether decreases in pausing over time serve a causal function for the development of larger pitch changes; (2) to determine the extent to which the growth of prosody during grades 1 and 2 is predictive of oral reading fluency in grade 3; (3) to determine whether the growth of reading prosody during grades 1 and 2 is predictive of comprehension skill in grade 3;

and (4) to determine the extent to which the development of reading prosody adds to our ability to account for reading fluency and comprehension outcomes beyond word reading speed and accuracy. Path model tests found evidence of a relationship between a decrease in the number of pauses during oral reading and the subsequent development of adult-like pitch contours. Furthermore, outcome model tests indicated that while aspects of both pause and pitch variables initially impacted oral reading fluency, only pitch contour emerged as a significant predictor of fluency once word reading speed and accuracy were taken into account. Finally, the cumulative effect of decreases in pausing and the initial pitch contour measurement predicted comprehension skill.

INDEX WORDS: Prosody, Oral Reading Fluency, Reading Development, Syntax

THE DEVELOPMENT OF PROSODIC TEXT READING AS A DIMENSION OF ORAL READING FLUENCY IN EARLY ELEMENTARY SCHOOL CHILDREN

by

JUSTIN MILLER

B.A., Rutgers University, 2003

M.A., University of Georgia, 2005

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSPHY

ATHENS, GEORGIA

2008

THE DEVELOPMENT OF PROSODIC TEXT READING AS A DIMENSION OF

ORAL READING FLUENCY IN EARLY ELEMENTARY SCHOOL CHILDREN

by

JUSTIN MILLER

Major Professors: Paula J. Schwanenflugel
A. Michele Lease

Committee: Jonathan M. Campbell
Shawn M. Glynn

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2008

DEDICATION

To my family and friends for their love and support.

ACKNOWLEDGEMENTS

This research was supported in part by the Interagency Education Research Initiative, a program of research jointly managed by the National Science Foundation, the Institute of Education Sciences in the U.S. Department of Education, and the National Institute of Child Health and Human Development in the National Institutes of Health (NIH Grant No. 7 R01 HD040746-06). I would like to acknowledge Patricia Foels, Elizabeth Meisinger, and Jennifer Sieczko for their help in data collection, Dr. Shawn Hendricks for his technical advice and assistance, and Drs. Jonathan M. Campbell, Arnold L. Glass, Shawn M. Glynn, and Melanie R. Kuhn for their input in this research. Most of all I would like to thank my advisors, Drs. Paula J. Schwanenflugel and A. Michele Lease, for everything that they have done for me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Since the publication of the report of the National Reading Panel (2000), and more recently the National Assessment of Educational Progress (NAEP) Special Study of Oral Reading (2005), fluency research has experienced something of a revival of interest. Although recognition that the development of reading fluency holds substantial importance is hardly a novel concept within the disciplines of psychology and education, fluency has often been described as the “most neglected” reading skill (Allington, 1983). The reasons for this, however, may have more to do with historical trends in the field of psychology and the continuously evolving conceptions of fluency (i.e., the particular subskills or qualities included in this aspect of reading performance), rather than the relative disinterest among researchers. In fact the foundations of research in reading fluency can be traced back to the work of experimental psychologists during the late 19^{th} century. Admittedly, much of the research conducted in the late 1800s and early 1900s characterized fluency primarily as “the immediate result of word recognition proficiency” and often neglected other components that are considered essential aspects of fluency today (National Reading Panel, 2000). Interest in fluency and other psychological processes faded, albeit somewhat involuntarily, during the period from approximately 1910 to the middle 1950s, while behaviorism dominated American social sciences and education. Although this resulted in the temporary suspension of research productivity and progress in the field, the onset of the cognitive revolution in psychology during the

1960s, and particularly the 1970s, revived research efforts in reading processes and sparked perhaps the most active era in the study of reading fluency. This philosophical shift recast research and theory on reading development and, consequently, positioned skill acquisition, the development of expertise, and automatic processing as the primary foci for decades to follow. Whereas the tradition of scholarship in reading fluency continues today, current research has taken on a heightened practical importance with recent revelations concerning the development of reading fluency in children and its role for academic success.

Importance and Associated Outcomes of Fluent Reading

The importance of fluency in the field of reading is derived largely from (1) the recognized correlation between developed expertise in reading and comprehension skill, and (2) the relationship between fluency and overall educational achievement. A large study examining the status of fluency achievement in American education conducted by Pinnell, Pikulski, Wixson, Campbell, Gough, and Beatty (1995), in association with the National Assessment of Educational Progress (NAEP), found that 44% of a nationally representative sample of 4^{th} grade students were disfluent with grade-level passages. Furthermore, Donahue, Voelkl, Campbell, and Mazzeo (1999) found that 38% of fourth grade students read below the “basic” level, which is defined as “partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.” These figures are particularly alarming given the additional finding of a significant relationship between oral reading fluency and reading comprehension, as measured by overall reading proficiency on the main NAEP assessment. A follow-up study by Daane, Campbell, Grigg, Goodman, and Oranje (2005) yielded similar results, as 39% of the nation's fourth graders were characterized as disfluent and oral reading fluency again shared a positive relationship with reading comprehension.

It follows that children who do not develop fluency early on in the schooling process will likely experience difficulty learning important material from texts introduced in later grades (Chall, Jacobs, & Baldwin, 1990; Lyon, 1997). This circumstance appears to be the direct result of the typical progression of school-based instructional methods. As school curriculum begins to shift and instructional emphasis in reading that was once based upon text information generally known to children is replaced by text information that is new, children are not only expected to learn independently from text but face significant risk of educational underachievement or failure should they lack the required skills to do so (Chall, 1996b). Indeed Rasinski, Padak, McKeon, Wilfong, Friedaur, and Heim (2005) found that reading fluency continues to be a significant variable in secondary students' reading and overall academic development. Accordingly, when the National Research Council (Snow, Burns, & Griffin, 1998) issued its report on the prevention of reading difficulties in young children not only did they recognize the importance of achieving fluency with various texts, but also recommended regular classroom assessment and effective instruction since "the ability to obtain meaning from print depends so strongly on fluency" (p. 323).

Current Status of Research in Reading Fluency

Despite a substantial, cumulative body of research examining a multitude of processes with suspected involvement in the development of fluent reading, and an understanding of the educational outcomes associated with the ability to read fluently, the construct of fluency has been criticized for lacking clear theoretical and definitional consensus in the research literature. Kame'enui and Simmons (2001), for example, characterized fluency as "a term so broad and unsatisfactory in meaning that little insight and understanding are gained beyond mere use of the term" (p. 204). Although theoretical disagreements may exist and definitions vary, the general principles underlying research in reading fluency appear to be consistent (some even contend that the aforementioned conflicts are more apparent than real). That fluency is a developmental process representing an outcome of sublexical and lexical processes and skills that are acquired during the elementary school years is seemingly unquestionable (Fuchs, Fuchs, Hosp, & Jenkins, 2001). That it is comprised of some combination of the elements of rate, accuracy, and "expression," for the purposes of comprehension is commonly agreed upon as well. Definitional inconsistencies, however, are often a reflection of advances in the field which involve mainly the inclusion of additional skills or qualities (e.g., prosodic, or "expressive," reading) that were not considered in traditional conceptions of fluency. Furthermore, an examination of early research successfully illustrates the establishment of enduring theoretical concepts that remain fundamental to the study of fluency at present. This certainly does not suggest that a single consolidated focus is currently guiding research efforts, but rather refutes the characterization of fluency as an overly fractured and disorganized construct. Criticisms, however, are not entirely misplaced. General agreements notwithstanding, the exact cognitive mechanisms and processes that index fluency, and the manner in which they do so, appear to be unsettled theoretically and experimentally (Kame'enui & Simmons, 2001; National Reading Panel, 2000; Stanovich 2000).

Early Research and Definitions of Fluency

As Pikulski and Chard (2005) point out, defining reading fluency as a construct is central to making important decisions about the ways in which fluency is studied. Furthermore, the process of doing so creates an additional understanding of how the study of fluency has evolved. A historical review of the term and its application, conducted by the National Reading Panel (2000), traced the “changing concepts of fluency” from “high-speed word recognition” fluency to processes that extend “beyond word recognition” to “comprehension processes as well” (chap.3, p.6; Kame’enui & Simmons, 2001). A more thorough treatment, however, would expand beyond the National Reading Panel analysis and include selected examples of early research, review of essential theoretical principles, and a survey of definitions that have been presented over time.

Indeed the foundation of fluency research emphasized a contextually-based speed component achieved as an outcome of automatic processing (Cattell, 1886). Specifically, Cattell suggested that “when the words make sentences and the letters words, not only do the processes of seeing and naming overlap, but by one mental effort the subject can recognize a whole group of words or letters, and by one will-act choose the motions to be made in naming them, so that the rate at which the words and letters are read is really only limited by the maximum rapidity at which the speech-organs can be moved” (p. 64). Additional studies have validated Cattell’s findings that speed of processing print is increased by orthographic, lexical, and semantic-syntactic information in both adults (see Barron & Pittenger, 1974; Doehring, 1976; Eichelman, 1970; Forster & Chambers, 1973; Reicher, 1969; Wheeler, 1970) and children (see Biemiller & Levin, 1968; Gibson, Osser, & Pick, 1963; Gibson, Barron, & Garber, 1972; Levin & Biemiller, 1968; Thomas, 1968). Sole emphasis on speed, however, is not necessarily a misguided focus since few would argue that oral reading, even if perfectly accurate, could be considered fluent if it were not rendered rapidly. Moreover, automatic, or speeded, performance presumably permits higher-order processes such as comprehension or prosodic reading to take place.

G. Stanley Hall focused on the underlying processes of skilled reading as well and formally introduced the concept of “automatization of function,” a term that he used to refer to “the process whereby well practiced events are run off with less and less conscious control.” Hall (1911) spoke of this automaticity when he defined true reading as “occurring when the art has become so secondarily automatic that it can be forgotten and attention can be given solely to the subject matter. Its assimilation is true reading and all else is only the whir of the machinery and not the work it does” (p. 445). As a consequence of this advanced processing, or automatization of function, skilled readers are less dependent on processing the printed word and may instead focus on constructing meaning.

This tradition of theory was continued by Huey (1908/1968) with the added notion that fluent reading develops in stages, with characteristically unique processes occurring at the different stages. Inherent to this belief was the idea that fluent reading involved the gradual accumulation and synthesis of complex processes and skills formed through practice. Thus the development of reading fluency was dependent on increasing the rate of processing through repetitive practice as “repetition progressively frees the mind from attention to details, makes facile the total act, shortens the time, and reduces the extent to which consciousness must concern itself with the process” (p. 65).

Although early researchers are responsible for establishing much of the conceptual, and even experimental, groundwork, most discussions of fluency trace their modern theoretical foundations to the work of David LaBerge and S. Jay Samuels. The seminal article on automatic information processing in reading by LaBerge and Samuels (1974) represents the most comprehensive attempt at modeling the complex processes involved in reading skill acquisition and is considered the most influential and widely quoted of all the reading theories (Blanchard, Rottenberg, & Jones, 1989). Automaticity theory suggests that children who have developed automatic word reading skills rapidly process high-frequency words and decode new words quickly, allowing attention to be shifted to reading for meaning. Evidence for this theory is speed- and accuracy-linked improvement in higher-level aspects of reading such as comprehension (Gough, 1996; Miller & Schwanenflugel, 2006; Nicholson, 1999; Perfetti & Hogaboam, 1975; Schwanenflugel, Hamilton, Kuhn, Wisenbaker, and Stahl, 2004). This conceptualization of fluency has been the cornerstone of research in reading fluency for more than 4 decades and its influence is evident in numerous definitions given since that time.

According to Wolf and Katzir-Cohen (2001), throughout the period between LaBerge and Samuels' (1974) automaticity theory and Carver's (1991, 1997) rauding theory (a theory that focuses on the links between fluency and comprehension through emphasis on the different purposes of reading and their respective rates), fluent reading was commonly defined as "that level of reading competence at which textual material can be effortlessly, smoothly, and automatically understood" (Schreiber, 1980, p. 177). A survey of additional definitions revealed that many others incorporated the same general principles of automatic text reading skills for the purposes of comprehension as well. For example, *The Literacy Dictionary* defined fluency as “freedom from word identification problems that might hinder comprehension” (Harris & Hodges, 1995, p. 95). Logan (1997) characterized fluent reading as “speeded, seemingly effortless, autonomous, and achieved without much consciousness or awareness.” Similarly, Meyer and Felton (1999) defined fluency as “the ability to read connected text rapidly, smoothly, effortlessly, and automatically with little conscious attention to the mechanics of reading such as decoding” (p. 284).

Whereas the approach to fluency described above provides a more complete enumeration of the behaviors involved in skilled reading and further incorporates the end of goal of comprehension, difficulties arise in validating such definitions and appropriately conceptualizing the underlying component structure of fluency (Wolf & Katzir-Cohen, 2001). Although some researchers support more simplistic definitions of fluency such as “rate and accuracy in oral reading” because they allow for relative ease in empirical validation (Shinn, Good, Knutson, Tilly, & Collins, 1992), others (e.g., Wolf & Katzir-Cohen, 2001) contend that such emphasis “ignores the multi-dimensionality of fluency” (p. 218). More recent definitions have expanded our understanding of fluency beyond word reading automaticity and comprehension, and included the additional element of prosodic or “expressive” oral reading. The National Reading Panel (2000), for example, defined fluency as the ability to “read text with speed, accuracy and proper expression” (chap. 3, p.1) and further included specific recognition that “fluency requires the rapid use of punctuation and the determination of where to place emphasis or where to pause to make sense of a text” (chap 3., p. 6). Alternatively, Pikulski and Chard (2005) proposed a synthesis of the National Reading Panel and The Literacy Dictionary definitions that reads as follows: “Reading fluency refers to efficient, effective word recognition skills that permit a reader to construct the meaning of text. Fluency is manifested in accurate, rapid, expressive oral reading and is applied during, and makes possible, silent reading comprehension” (p. 510). Still, Wolf and Katzir-Cohen (2001) provided perhaps the most comprehensive working definition of reading fluency which they derived from a combination of the developmental (Kame’enui, Simmons, Good, & Harn, 2001) and systems-analysis perspectives (Berninger, Abbott, Billingsley, & Nagy, 2001). The authors characterize the beginnings of fluency as “the product of the initial development of accuracy and the subsequent development of automaticity in underlying sublexical processes, lexical processes, and their integration in single-word reading and connected text” which include “perceptual, phonological, orthographic, and morphological processes at the letter, letter-pattern, and word levels, as well as semantic and syntactic processes at the word level and connected-text level” (p. 219). After it is fully developed, “reading fluency refers to a level of accuracy and rate where decoding is relatively effortless; where oral reading is smooth and accurate with correct prosody; and where attention can be allocated to comprehension” (p. 219).

Syntactic Structure and Prosodic Extraction in Speech and Reading

The belief that the production of prosody is a reflection of the syntactic structure of a given sentence is widely supported in the fields of linguistics and psychology (Ferreira, 1993). According to models of prosodic structure (Nespor & Vogel, 1987; Selkirk, 1986), syntax does in fact appear to influence the prosodic characteristics of a sentence but may do so indirectly. This indirect effect results because the syntactic structure of a sentence ultimately influences the sentence’s organization into prosodic segments. For example, Ferreira (1993) points out that in models of prosodic structure, a syntactic-phrase boundary virtually forces a phonological-phrase boundary which is the reason that these boundaries are often the focus of various prosodic features, including pitch fluctuation, pausing, and phrase-final lengthening. Although Ferreira’s (1993) model demonstrated that prosodic timing patterns show a general correspondence to syntactic structure, it also stressed that syntactic and prosodic structures may not necessarily be identical.

Whereas the majority of research and theory on prosody has focused on speech rather than reading, Koriat, Greenberg, and Kreiner (2002) propose that prosody applied during text reading reflects the structural framework established for phrases/sentences as well. In fact, reading prosody may be more closely aligned with grammatical structure than speech prosody (Goldman-Eisler, 1972). Goldman-Eisler found that during spontaneous speech production less than one-third of breathing pauses occurred at clause boundaries; however, breath pauses almost always occur at such boundaries during oral reading. Moreover, Koriat et al. (2002) found that prosodic reading appears to be derived primarily from syntactic structure prior to, and independent of, semantic information. Although the authors recognize that semantics may play a minor role in prosody, they suggest that reading may be more similar to speech comprehension than to speech production. That is, individuals (even at very early perceptual stages) are responsive to structural cues in speech that are necessary for organizing information prior to the construction of meaning (Koriat et al., 2002). Thus, the authors speculate that prosody may serve an important representative function of early structural analysis and that this prosody is crucial for retaining information in working memory so that it may then be processed.

Prosodic Structures in Infant's Understanding and Use of Language

Sensitivity to the use of prosodic features in speech is especially evident in young children (Dowhower, 1991; Schreiber, 1987; Schreiber & Read, 1980). Research suggests that infants not only use prosody as a primary cue to the syntactic structure of their language, but that their babbling mimics the prosodic characteristics inherent in their primary language as well (Kuhn & Stahl, 2003). Furthermore, Read and Schreiber (1982) and Schreiber (1987) demonstrated an additional finding which suggested that children appear to be more reliant on prosodic elements in oral language for determining meaning than are adults.

Theoretical accounts of the prosodic structure of spoken language (Nespor & Vogel, 1986; Selkirk 1983) propose a "hierarchy of elements ranging from morae (the minimal unit of sound used in phonology that determines syllable weight) and syllables to intonational phrases and utterances" where elements in the hierarchy above the word-level are derived from syntactic structure and serve as the domains within which various phonological rules apply (Gout, Christophe, & Morgan, 2004, p. 550). In other words, prosodic words constitute phonological phrases (whose boundaries coincide with syntactic phrase boundaries), which in turn constitute intonational phrases that are most often whole clauses (Gout et al., 2004). For purposes of clarity, the term *phonological phrase* refers to any kind of prosodic level above the prosodic word (one or more prosodic words, under the general rule of two to three prosodic words), whereas the *intonational phrase* consists of one or more phonological phrases (under the general rule of two to three phonological phrases). Hirsch-Pasek, Nelson, Jusczyk, Cassidy, Druss, and Kennedy (1987) suggested that young infants perceive intonational phrase boundaries. Specifically, they observed that infants listened longer to stimulus sets when pauses were inserted at intonational phrase boundaries as opposed to when inserted at other, non-boundary, points. Furthermore, numerous studies have demonstrated that infants are sensitive to prosodic cues and syntactic boundaries that correspond with phonological phrase boundaries (Christophe, Dupoux, Bertoncini, & Mehler, 1994; Christophe, Mehler, & Sebastian-Galles, 2001; Gerken, Jusczyk, & Mandel, 1994; Jusczyk et al., 1992). Gout et al. (2004) conducted a series of studies the results of which indicated that infants as young as 10 months old may use cues to phonological phrase boundaries to segment connected speech as well.

According to Nazzi, Dilley, Jusczyk, Shattuck-Hufnagel, and Jusczyk (2005), word segmentation, or the extraction of the sound patterns of words from the speech signal, is a critical step in infant speech processing. Recent work examining the factors involved in early word segmentation highlighted the importance of prosodic cues, particularly lexical stress at the onset of word segmentation, in identifying word boundaries. The early acquisition and use of prosodic information is supported by studies suggesting that infants between the ages of 6 and 9 months show a preference for “words with the predominant English strong-weak (SW) stress pattern (e.g., *porter*) over less frequent weak-strong (WS) words (e.g., *report*)” (Nazzi et al., 2005, p. 280; see also Jusczyk, Cutler, & Redanz, 1993; Turk, Jusczyk & Gerken, 1995). In fact additional studies (using various adaptations of the *head-turn preference procedure*) have demonstrated that infants begin segmenting nouns with the SW pattern at approximately

7.5 months, whereas they start segmenting WS nouns at 10.5 months (Jusczyk, Houston, & Newsome, 1996b; see also Echols, Crowhurst, & Childers, 1997; Houston, Santelmann, & Jusczyk, 2004; Johnson & Jusczyk, 2001; Morgan & Saffran, 1995).

Nazzi et al. (2005), conducted similar studies examining verb stress patterns and found a notable developmental lag in segmentation compared to nouns (13.5 months for strong-weak consonant- or vowel-initial verbs and for weak-strong consonant-initial verbs; and 16.5 months for weak-strong vowel-initial verbs). The authors suggested that pitch accent and phrasal boundary distributions could account for some of the performance differences between the studies. Specifically, prosodic analyses revealed that nouns were more likely to be clearly followed by a phrasal boundary and more likely to be preceded by a syllable bearing a pitch accent (Nazzi et al.). Accordingly, clear demarcation of noun stimuli as used in previous studies could account for infants' relative ease in segmentation compared with verbs. Nevertheless, these studies support the conclusion that infants are perceptually sensitive to prosodic cues in processing speech.

Given the evidence demonstrating that children's understanding of oral language is to some extent dependent upon the use of prosodic features, one could reasonably assume that prosody is an important determining factor in children's ability to derive appropriate meaning from text as well (Allington, 1983; Dowhower, 1991; Kuhn & Stahl, 2003; Schreiber, 1991). In fact, appropriate phrasing, intonation, and stress are all considered to be indicators of fluent reading (Chomsky, 1978; Rasinski, 1990b; Samuels, Schermer, & Reinking, 1992) and are further thought to reflect the otherwise invisible process of comprehension (Kuhn & Stahl, 2003). Specifically, Kuhn and Stahl assert that the ability to group text into syntactically appropriate phrases signifies that a reader has an understanding of what is being read.

Contribution of Prosody in Reading Comprehension

Whereas prosodic reading is widely considered to be a hallmark of the achievement of reading fluency (Dowhower, 1991; Kuhn & Stahl, 2003; Schwanenflugel et al., 2004), the link between prosody and other aspects of the reading process remains unclear. Although automaticity theory accounts for the accurate and effortless decoding that is characteristic of fluent reading, and further provides an explanation for potential increases in comprehension skill, it does not explicitly address the role of prosody in the reading process. Kuhn and Stahl (see also National Reading Panel, 2000) suggest that prosodic text reading is a necessary condition, beyond automatic individual word decoding, for adequate comprehension to occur. According to the authors, this reasoning is based on the theoretical proposition that the development of reading prosody may assist comprehension because prosodic reading is indicative of the ability to segment text according to major syntactic/semantic elements. Support for this argument is evident, as research demonstrates that comprehension may be related to skill in syntactic phrasing (Young & Bowers, 1995) and comprehension improves when children are provided with information about syntactic and semantic boundaries (Cromer, 1970; O'Shea & Sindelar, 1983). Consequently, prosody might serve an important function in the process of reading by providing the bracketing of key syntactic and semantic boundaries that not only signal the ability to group text into meaningful phrase units, but indicates that the reader has an understanding of what is being read as well.

That prosody may provide an important linguistic link between fluency and comprehension is a relatively well-established concept in the research literature; however, evidence is somewhat inconclusive. Chafe (1988) suggested that to read a sentence with intonation, one must assign syntactic roles to the words in the sentence. The assignment of syntactic roles is a key component of microprocessing, or the mental parsing of a text into hierarchically ordered propositions (Kintsch, 1998). Schreiber (1987) suggested that the explicit presence of prosodic cues might be one crucial difference between speech and reading, and is one of the reasons that speech is easier to understand. Assuming that prosodic cues serve an important signaling function for children in their processing of spoken language (Morgan, 1996), it follows that the absence of such cues in print may partially account for the difficulty many children have in parsing written text. However, Schreiber reported that evidence supporting a link between prosody and syntactic processing is weak, with some studies finding links between the use of prosodic features and syntactic comprehension and others failing to find such an effect.

Punctuation as a Cue to Prosodic Interpretation

Punctuation may serve as the visual cue to syntax-related prosody. Recently, Steinhauer (2003) suggested that overt prosody in spoken language and implicit prosody cued by punctuation during reading may have strong influences on sentence comprehension by guiding syntactic parsing. Steinhauer conducted experiments examining the processing of commas in silent reading to determine primarily whether commas served as orthographic triggers for covert, or subvocal, prosodic phrasing. In these event-related brain potential studies, speech boundaries and commas reliably elicited a similar online brain response, termed the *Closure Positive Shift* (CPS). According to Steinhauer, this finding supported a direct correspondence between punctuation and implicit prosody, pointing to a common mechanism that allows punctuation in written communication to take on the functions that prosody does for speech.

In contrast, Chafe (1988) argued that although punctuation is intended to capture major aspects of prosodic intent, prosodic features are not always well dictated by text punctuation. For example, he noted that grammar rules that govern the placement of phrase-final commas between words in a series may dictate pauses in sentences like *He came, he schmoozed, and he dazzled*, but not in sentences such as *Arnie wanted the one with the red, white, and blue sprinkles*. Question marks also exhibit prosodic uncertainty, seeming to dictate a final pitch rise for the end of yes-no questions *(e.g., Did Melanie go?),* but not for wh-questions *(e.g., Where did Melanie go?).* Moreover, spoken language consists of shorter speech segments (about 5 or 6 words) before pausing than would be dictated by written punctuation, particularly for lengthy sentences. Consequently, oral readers will introduce their own prosodic boundaries not signaled in the text at all. Thus, because oral readers must abstract prosodic features to a great extent while reading aloud, one of the tasks children have in learning how to read aloud is to learn the limits of punctuation as a cue to the underlying prosodic structure of the text.

Besides learning the limits of punctuation, elementary school children are still developing their understanding of prosody. Bates (1976), for example, found that prosodic stress patterns are processed poorly by children as old as 8 years of age. Furthermore, Cutler and Swinney (1987) found that even 9- and 10- year-olds are not quite at adult levels in understanding the function of some contextual prosodic features. Thus, it is possible that children just learning to read may not be able to make full use of the prosody engaged by oral reading.

Prosodic Measurement and Studies Examining the Role of Prosody in the Reading Process

According to Dowhower (1991), scholars have identified at least six distinct prosodic indicators related to expressive reading: (a) pausal intrusions; (b) length of phrases; (c) appropriateness of phrases; (d) phrase-final lengthening; (e) terminal intonation contours; and (f) stress. Taken together, these features are classified as suprasegmental because they extend over more than one speech sound and contribute to meaning. As previously noted, appropriate use of such markers signifies a reader’s ability to apply syntactic knowledge to text and further demonstrates the ability to produce the essential features of expressive oral language during reading, while preserving accuracy and speed.

Currently, the majority of available studies examining the development of oral reading prosody have focused on measures such as descriptive ratings of specific prosodic features (Bear, 1992; Clay & Imlach, 1971). Clay and Imlach, for example, used a rater to analyze separately the pausing, pitch, and stress present in the oral readings of seven-year-old children and found that children who made few and short pauses were the best readers according to objective assessments of skill. In addition, more highly skilled readers completed declarative sentences with a declination in pitch. However, because statistical analyses were not carried out, it is unclear whether the skill differences were reliable across children. Further, ratings can be unreliable across raters for some prosodic features because it is difficult for listeners to disentangle prosodic from decoding issues (Bear, 1992).

Although research employing direct measurement of reading prosody is surprisingly sparse, several studies have attempted to examine prosody directly through spectrographic analysis. For example, Herman (1985) counted the presence of speech pauses of eight remedial fourth to sixth grade children as they carried out repeated oral readings of a moderately difficult text and found that the number of pauses not dictated by punctuation dropped considerably as a result of repeated reading. As noted earlier, punctuation may be only a very rough indicator of where pauses are appropriate and the number of participants used in this study was quite small. Dowhower (1987) examined the effect of repeated reading on oral reading prosody in 2nd grade children who read accurately but in a slow and word-by-word manner. Students' audio-taped oral readings were analyzed spectrographically to determine the duration of each word, the length of pauses between words, and the fundamental frequencies (F_0) for subject-final and sentence-final words. After repeated practice, children showed significant improvements in prosodic reading in terms of decreased inappropriate pausing within words or major syntactic units, increased sentence-final vowel lengthening (a prosodic feature marking the end of a major syntactic unit; Cooper & Paccia-Cooper, 1980), and a greater F_0 declination occurring at the last syllable of a declarative sentence.

Schwanenflugel, Hamilton, Kuhn, Wisenbaker, and Stahl (2004) examined the role of reading prosody in order to (1) characterize the development of prosodic reading as a function of reading skill, and (2) test the model that prosody may serve as a partial mediator between decoding and comprehension skills. Spectrographic analysis was conducted on the oral readings of a large sample of second and third grade children with the purpose of investigating the following five prosodic features: (1) inter-sentential pause length; (2) inter-sentential pause length variability; (3) intra-sentential pause length; (4) child-adult F_0 sentence profile match (based upon adult comparison sample data); and (5) the sentence-final F_0 declination. According to the authors, skilled readers (as determined by standardized assessment of word reading efficiency) were found to make shorter pauses both within, and between, sentences, with minimal variability. Further, good oral readers ended declarative sentences with discernable and relatively large pitch declinations, as noted in Clay and Imlach (1971) and Dowhower (1987). Skilled oral readers matched adults in their overall prosodic contours as well. However, although the authors found a clear connection between prosodic reading and word reading efficiency, reading prosody itself added little to predicting comprehension skills beyond that accounted for by word reading efficiency alone. Consequently, they argued that reading prosody should be viewed as an indicator of the emergence of automatic word reading skills.

Whereas Schwanenflugel et al. (2004) successfully characterized prosodic text reading as a function of word reading efficiency in young children, there are a number of possibilities for why the authors failed to find a significant relationship between reading prosody and reading comprehension skill. First, that study focused on children’s reading of a simple seven sentence pre-primer passage which consisted mainly of declarative sentences. As a result, this text may have lacked the structural complexity and breadth of prosodic features necessary to establish the relationship with reading comprehension skills. In other words, simple passages may not encourage children to mark prosody in a way that relates to comprehension. The falling fundamental frequency found for declarative sentences may serve as the "default" for sentence prosody, and may not represent marked or contrastive prosody. It may be that, as texts become more complex, children draw on their prosodic resources in a way that is more reflective of comprehension processes. There is also a minor possibility that developmental change, resulting from use of a sample that consisted of a mixture of second and third grade, obscured potential prosody-reading skill relations, given that speech prosody is also under development to some extent at this age.

Miller and Schwanenflugel (2006) continued to examine the role of prosody in the overall reading process and expanded their efforts to focus on the relationship between the prosodic reading of syntactically complex sentences, reading speed and accuracy, and comprehension skill. As a follow-up to the Schwanenflugel et al. (2004) study, the authors addressed a number of limitations identified in earlier studies, the most important advancement, however, involved the use of a text that included redundant observations of a more complete set of grammatical features than identified in previous research on reading prosody.

The design of the Miller and Schwanenflugel study targeted the following features which Chafe (1988) and Cooper and Paccia-Cooper (1980) suggested might require a distinct prosodic reading in adults: basic declaratives, basic quotatives, wh-questions, yes-no questions, complex adjectival phrases, and phrase-final commas. For example, intuitively it may seem that questions should be marked with a rising pitch, but Chafe (1988) indicated that this is true for some question types only. Similarly, intuition might dictate that commas should be marked with a pause, but Chafe suggested that not all commas need to be marked. Some, like phrase-final commas (such as *To avoid being run over, Paul jumped back*), might need to be marked while others, such as commas in complex adjective phrases (such as the *large, striped, yellow bus)* might not be. Finally, quotatives (such as "*We oughta go*," *said Freddy*) were another type of sentence that seemed to call on prosodic marking.

Therefore, prior to any investigation concerning prosody's position in the reading process of children, the authors first used an adult sample to discern which syntactic features are marked prosodically and which are not. In agreement with previous studies, adult readers reliably marked basic declarative sentences with a pitch declination. Although it was suggested that basic quotatives may require a pause following a quote (Cooper & Paccia-Cooper, 1980), adult readers did not pause following basic quotatives. According to Chafe (1988), wh-question types may not require an upswing in pitch; in support of this, the authors found that adults did not show a uniform treatment of this structure, with many adults showing a moderate to large pitch rise and others electing to end these questions with a pitch decline. Further, Chafe suggested that adults mark yes-no questions with an upswing in pitch. Miller and Schwanenflugel (2006) found that adult readers did in fact mark yes-no question types with a relatively large rise in pitch. Chafe also suggested differentiated reading of internal comma structures, specifying that pauses are not marked at commas in a series (e.g., *happy, playful, curious…*), but may be marked for phrase-final locations (e.g., *One afternoon, near a pond…*). However, it was found that adult readers generally did not pause at either structure. Overall, however, the findings concurred with Chafe's basic point that that punctuation does not seem to drive prosodic readings in adults.

Most importantly, Miller and Schwanenflugel (2006) found that skilled child readers seemed to be heading toward a prosodic rendering that was similar to adults. The findings presented regarding the prosody-reading skill relationships showed general similarity to those described in previous studies and added some new findings as well. Like Clay and Imlach (1971) and Schwanenflugel et al. (2004), the authors found that good readers made short pauses; however, new information regarding children's treatment of internal punctuation extended our current understanding and indicated that good readers kept pauses short at internal commas across a variety of sentence types (i.e., basic declaratives, basic quotatives, and yes-no questions), but less skilled readers did not. Similarly, agreement was also evident with Clay and Imlach, Dowhower (1987), and Schwanenflugel et al. that skilled readers ended declarative sentences with a marked declination in pitch. However, Miller and Schwanenflugel added the observation that skilled readers show comparatively large pitch rises following yes-no questions. Thus, the hypothesis that reading prosody is a feature that emerges once children have acquired quick, accurate, and automatic word- and text-level reading skills (LaBerge & Samuels, 1974; Perfetti, 1985) showed strong support, as reading skill was related to short and more adult-like pause structures, large declinations at the ends of declaratives and larger pitch rises following yes-no questions.

Similar to the Schwanenflugel et al. (2004) study, Miller and Schwanenflugel (2006) also tested an additional hypothesis that, once prosodic reading was established, this prosody of syntactically complex sentences might make a unique contribution to comprehension skill beyond that accounted for by quick and accurate text reading alone. As mentioned previously, this view was based on previous research that suggested prosodic reading may provide important syntactic and semantic feedback to the reader which may ultimately assist comprehension (Cromer, 1970; Kuhn & Stahl, 2003; O'Shea and Sindelar, 1983). The authors found support for this hypothesis, but only for specific pitch features. That is, children who showed large declinations at the ends of basic declarative sentences and larger pitch rises following yes-no questions tended to be those whose comprehension skills were higher. Pausing, however, was unrelated to comprehension skills beyond that accounted for by rapid and accurate text reading.

Taken together, the Schwanenflugel et al. (2004) and Miller and Schwanenflugel (2006) studies were successful in characterizing prosodic reading as a function of reading skill and provided better evidence for the role of prosody on reading comprehension compared to previous research. Although neither study found an independent effect of pause structures on reading comprehension once reading speed and accuracy were taken into account, their findings for pitch change were somewhat mixed. That is, while Schwanenflugel et al. found no effect for sentence-final F_0 declinations for declarative sentences on reading comprehension, they did, however, note a small, but significant effect of child-adult F_0 contour match. Children whose general pitch contour was similar to that of the average adult tended to have higher reading comprehension skills. However, they reached the ultimate conclusion that prosody, on the whole, did not add much to the ability to predict children's reading comprehension beyond that accounted for by individual word decoding skills. Alternatively, Miller and Schwanenflugel demonstrated a somewhat more general effect of sentence-final pitch change on reading comprehension skill. The reason for this is believed to be the result of a focus on syntactically complex sentences and the use of a passage which was a closer match to children's overall reading skill level. Regardless, the findings presented suggest that different aspects of prosody may be distinctly related to different aspects of the reading process. Long pauses may signal general difficulties with decoding skills. By contrast, large sentence-final pitch changes (where appropriate) appear to be an important variable in the prediction of reading comprehension skill.

Purpose of the Present Study

Current research considers fluency to be comprised of (1) accuracy in decoding, (2) automaticity in word recognition, and (3) the appropriate use of prosodic features such as stress, pitch, duration, and appropriate text phrasing, all of which are assumed to facilitate text comprehension (Kuhn & Stahl, 2003; Wolf & Katzir-Cohen, 2001). While there is a rich literature about the systematic development of automatic word recognition skills (Ehri, 1995; 1998) and the contribution of accurate word recognition to reading comprehension (Johns, 1993; Stanovich, 2000), relatively little is known about prosody. According to Chall’s (1996b) stages of reading development, the development of prosodic text reading occurs in the second of six proposed stages called confirmation and fluency, or “ungluing from print” (p. 18). During this stage readers “confirm” what is already known to develop their fluency and, having established accurate decoding skills in the previous stage, must now develop automaticity with text (Chall, 1996b; Kuhn & Stahl, 2003). Presumably, as their reading becomes increasingly fluid, children develop the ability to represent what is read in ways that imitate the tonal and rhythmic aspects of conversational speech. To do this requires use of prosodic features that include appropriate phrasing, pause structures, stress, rise and fall patterns, and general expressiveness. In a sense, then, this model supports a multi-dimensional view of fluency and suggests that the development of accurate decoding and automatic word recognition in connected text creates the conditions necessary for prosodic reading to occur. While research suggests that skilled readers are more likely to read prosodically (Clay & Imlach, 1971; Dowhower, 1987; Herman, 1985; Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004) and that prosody may make additional contributions to comprehension skill beyond those accounted for by quick and accurate word reading skills alone (Cromer, 1970; Kuhn & Stahl, 2003; Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004; O'Shea & Sindelar, 1983; Young & Bowers, 1995), the full range of prosodic features that might be implicated remains unclear. Furthermore, at present there are no systematic studies examining the development of prosodic reading independently, or in relation to the acquisition of other skills involved in fluent reading. According to Chall's model, children are expected to develop automaticity and ultimately prosodic text reading skills during the confirmation and fluency stage which spans from the end of first grade to third grade. However, the lack of empirical evidence to support this assertion represents a critical theoretical gap in the achievement of skilled reading.

The purpose of this study was to examine the development of prosodic, or expressive, text reading during grades 1 through 3. Given that prosody is widely considered one of the defining characteristics of fluent reading and believed to develop as a result of the successive acquisition of word- and text-level automaticity during the early elementary school years, it is of critical theoretical, and potentially instructional, importance to establish its place in the reading process. A longitudinal analysis is necessary to determine the growth trajectory of prosodic reading throughout the years of primary reading development and makes possible comparisons with skill development in other areas as well. A “long-term” longitudinal design is considered a more appropriate method for the current investigation not only because it offers several advantages over cross-sectional research (e.g., information about onset, continuity, prediction, and within-individual or subgroup change), but given the relatively complex nature of the construct to be studied as well. Specifically, a longitudinal study of prosodic growth offers the benefits of establishing more clearly the emergence of oral reading prosody in early elementary school readers and also allows for the observation of how the development of prosodic reading proceeds during the process of skilled reading acquisition. In this way, creating a context for prosodic reading in relation to the development of other skills (e.g., word and text reading automaticity, comprehension skill) as they occur within an individual or groups of individuals allows one to make predictions and formulate/confirm hypotheses about the process of learning to segment and mark text. This ultimately permits a more complete understanding of the role of prosody in reading theory. Although longitudinal research may be costly and time-intensive, and further may be subject to considerable sample attrition, it is particularly conducive to the observation of specific growth trends as they unfold in real-time. Neither cross-sectional nor two-wave studies provide a sufficient basis for studying development, particularly with regard to the subtle changes that may be present in reading development and evident in prosodic growth.

Given that we were interested in examining the role of prosody in the overall reading process, children were administered a battery of assessments that included measures of word reading efficiency, oral reading fluency, and reading comprehension (in addition to prosodic measurements). Data were collected on five separate occasions beginning during the spring of first grade and ending during the spring of the third grade school year. The current study proceeded in two analytic phases. First, it was necessary to determine which prosodic features might be relevant for a developmental study of reading prosody. Although a variety of syntax- and punctuation-based prosodic features have been discussed in previous research on reading and/or speech prosody as markers that evidence distinct prosodic interpretations, it is unclear which of these features display stable/logical patterns of change throughout the years of primary reading acquisition. In the first phase, a small-scale exploratory longitudinal analysis (using a portion of the total sample) was carried out to examine the full scope of features with potential relevance to the developmental study of reading prosody and to determine which features would be targeted for more thorough investigation in the larger longitudinal study. Once a set of target features was selected based on the findings of the preliminary analyses, we then proceeded with the full, large-scale longitudinal study of the development of prosodic text reading (hereafter referred to as Phase 2). The purpose of Phase 2 was three-fold and concerned the following: (1) an examination of the manner in which prosodic features develop in the process of skilled reading; (2) whether, and the extent to which, the early development of prosodic reading on simple passages predicts the development of fluent reading and comprehension as outcomes; and (3) an analysis of how decoding skills might impact the relationships between the early development of prosodic text reading and subsequent reading fluency and reading comprehension skills.

CHAPTER 2

METHOD

Participants

Participants were 92 first-grade students (40% male, 60% female; mean age = 7 years, 2 months; SD = 4 months; range = 6 years, 4 months – 8 years, 1 month) who were part of a larger study of the development of reading fluency. The students were enrolled in one of five schools in northeast Georgia (2 high-poverty public schools, 2 rural public schools, and 1 private parochial school). Only children who were not currently receiving special services for English language learners were included in the study. Six subjects (separate from the 92 noted above) were excluded a priori because they were unable to read the target passage at a level from which meaningful prosodic measurements could be obtained. An additional 8 subjects (separate from the 92 noted above) were removed because their recordings were of insufficient quality to conduct prosodic analysis. Approximately 63% of the children were African-American, 20% European-American, 13% Hispanic-American, and 4% of unknown ethnicity. Children enrolled in the public education system came from schools in which approximately 72% of the students qualified for free or reduced lunch programs.

In addition, 34 adults from the children’s communities provided oral reading samples. These recordings were collected for the purpose of obtaining baseline prosodic measurements for use as a point of comparison with child reading prosody. Adults were recruited from schools, neighborhood restaurants, stores, and other public venues in proximity to the children's schools. Balanced numbers of middle- and working-class, male and female adults were sampled. Adult subjects were recruited by asking if they felt comfortable reading a children's passage aloud. They were paid $10 for their participation.

General Reading Assessments and Procedures

As part of the larger study of the development of reading fluency, all subjects were administered a battery of reading assessments which included formal measures of word reading efficiency, oral reading fluency, and reading comprehension. Initial measurements were taken during the spring of the children's first grade school year with follow-up assessments occurring at four additional time points: during the fall of second grade (mean age = 7 years, 7 months), winter of second grade (mean age = 7 years, 9 months), spring of second grade (mean age = 8 years, 2 months), and a final measurement in the spring of the third grade school year (mean age = 9 years, 1 month). Assessments were appropriately counterbalanced, such that half the subjects received the word reading efficiency and oral reading fluency measures in the first half of the battery and half received the reading comprehension assessment first. Several additional measures were administered as well; however, they were not relevant in the context of the present study because they were completed on a different day than the measures reported here. Data collection assistants were trained on administration and scoring procedures to the standard of 100 percent agreement with the lead assistant on all reading assessments immediately prior to collection at each data collection wave. An illustration of the overall design and assessments administered at each time point can be found in Figure 2.1.

Word reading efficiency assessment. To obtain an independent estimate of word reading efficiency, children were administered the Test of Word Reading Efficiency (TOWRE), Sight Word Efficiency and Phonemic Decoding Efficiency subtests (Torgesen, Wagner, & Rashotte, 1999). The Sight Word Efficiency (SWE) subtest assesses the number of real words correctly read from a list within 45 seconds, whereas the Phonemic Decoding Efficiency (PDE) subtest measures the number of pronounceable phonetically regular nonwords that can be accurately decoded within 45 seconds. Forms A and B were administered alternating across time points. Children were initially assessed during the spring of first grade using the TOWRE-Form A which was repeated during the subsequent fall of second grade time point. Form B was administered during the winter of second grade, Form A during the spring of second grade, and finally Form B during the spring of third grade. Thus the test administrations yielded an overall A-A-B-A-B pattern (with summer vacation occurring between the first two administrations). Concurrent validity estimates reported in the test manual have a median of .91 in Grades 1 through 3. Alternate form reliabilities have a median score of .97 in Grades 1 through 3. Raw scores from each subtest were used as indicators of word reading and decoding efficiency.

Oral reading fluency assessment. The Gray Oral Reading Tests, Third Edition—Form A was administered during the initial spring of first grade assessment to obtain an estimate of skill in reading connected text. Subsequent measurements made use of the Gray Oral Reading Tests, Fourth Edition—Forms A and B that were alternated across time points and mimicked the pattern described above. Specifically, children were administered the Gray Oral Reading Tests, Fourth Edition—Form A during the fall of second grade, Form B during the winter of second grade, Form A during the spring of second grade, and Form B again during the spring of third grade assessment. At each time point, children were presented with a series of passages to read aloud and were scored on the rate and accuracy of their reading. Rate and accuracy scores were combined to yield a standard fluency score for each passage read. The sum of the individual passage fluency scores was used as an indicator of connected text reading ability.

Reading comprehension assessment. The Reading Comprehension subtest of the Wechsler Individual Achievement Test (WIAT) was administered to obtain an independent measure of reading comprehension skill. This subtest consists of a series of printed passages, each of which increases in complexity and is followed by an orally presented question. The subtest contains both literal and inferential comprehension question types. The children were instructed to read a passage, listen to the question presented by the examiner, and then provide an oral response in his, or her, own words. The test was discontinued once a child missed four consecutive items as directed by the test manual. The WIAT measures reading comprehension as children's ability to answer questions about the text, a skill which many teachers consider a key indicator of reading comprehension (Richardson, Anders, Tidwell, & Lloyd, 1991). The test manual reports validity estimates at third grade with a median of .78 with other reading comprehension tests and a median reliability estimates of .91 in this age range (WIAT, 1992). The raw score, determined by the number of individual questions answered correctly, served as an indicator of reading comprehension skill.

Reading Prosody Assessment and Procedures

Stimuli and procedures. The longitudinal study of the development of prosodic text reading was embedded within the context of the larger developmental study of oral reading fluency. To conduct this investigation of reading prosody it was necessary to select measurement points for which identical assessments were administered and consistent prosodic measurements could be carried out over time. Consequently, we selected 3 out of the original 5 time points based on the administration of the Gray Oral Reading Tests. Measurements began during the mid-spring (March-May) of first grade (time 1) and proceeded at roughly six month intervals with follow-up assessments during the mid-fall (September-November) of second grade (time 2) and the mid-spring (March-May) of second grade (time 3). Prosodic measurements conducted during the initial first grade assessment were taken from children's reading of the first passage of the Gray Oral Reading Tests, Third Edition-Form A (GORT-3; Wiederholt & Bryant, 1992). Subsequent measurements during the fall second grade time point and spring second grade time point made use of the same passage; however this passage was now the third in the updated Gray Oral Reading Tests, Fourth Edition-Form A (GORT-4; Wiederholt & Bryant, 2001). This passage was selected because it was highly decodable and allowed for the assessment of prosodic reading in the absence of numerous decoding errors. Moreover, because this passage appeared in both the GORT-3 and GORT-4, it allowed for consistency of the target passage across time points. Furthermore, the technical manual reported that this passage is appropriate entry-level material for children at the first and second grades. Readability analyses were conducted using the Flesch-Kincaid Grade Level Formula and the Spache Readability Index. Readability was computed and averaged across indices, yielding an estimated grade level of 1.97. Examiners provided the students with general directions to read the passage as quickly and as well as they could. In addition, the passage was introduced with the following instructions: “This story is about two people in a family. Read the story to find out what happens to them.” The passage was presented as formatted in the student booklet and shown below:

> A man got out of the car.
>
> He had a pretty box under his arm
>
> A little girl ran from the house to meet him.
>
> “Hello, Father,” she said.
>
> “Do you have a surprise for me?”
>
> Father said, “I have something for a good girl.”
>
> The girl laughed, “I am very good.”

The measurement design for the developmental study of reading prosody can be seen in Figure 2.2.

Apparatus. Oral reading recordings were obtained for the target GORT passage with the goal of acquiring high quality recordings suitable for prosodic analysis. These recordings were obtained using a variety of equipment, a Sony TCD-D100 digital audiotape (DAT) cassette recorder, a Sony ECM-717 Stereo Unidirectional Microphone, or a Dell Inspiron 5100 notebook computer, Sound Devices USBPre 1.5 Microphone Interface, and a Sony ECM-717 Stereo Unidirectional Microphone. USBPre 1.5 is a complete, portable hardware interface for computer-based digital recording. All of these means of recording have been used in previous research examining reading prosody (see Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004). A shareware version of

GoldWave digital audio editor was used to create individual *wav* files. Noise reduction procedures were utilized to filter background interference. Prosodic analysis of these recordings was conducted using *Praat v*.4.3.07. *Praat* is a comprehensive speech software package designed to analyze, synthesize, and manipulate digital speech data (Boersma & Weenink, 2004).

Procedures for Selecting Developmentally Relevant Prosodic Features

Although theory suggests that prosodic reading develops as text reading skills become increasingly automatic, the exact features that index the developmental nature of prosodic text reading and the manner in which they do so has not yet been determined. Therefore, it was particularly necessary to include, and survey, the full scope of known features with potential developmental importance in order to identify those prosodic structures that would be targeted in the large-scale study. To this end, a preliminary exploratory longitudinal analysis was conducted with a small sample of children and a wide assortment of prosodic features (both syntax- and punctuation-based) that have been identified or implicated in previous research on reading and/or speech prosody as markers that evidence distinct prosodic interpretations. Selection decisions were based on (1) the extent to which a particular feature was one that displayed a pattern of change over time, and (2) the extent to which a particular feature demonstrated a distinct prosodic reading that was evident over the course of development and consistent with the suggested "target" reading as determined in previous research (see Clay & Imlach, 1971; Dowhower, 1987; Miller & Schwanenflugel, 2006; Schwanenflugel et. al., 2004). If, however, it was clear within this smaller sample that specific prosodic features failed to show clear developmental trajectories (in other words, those for which an erratic or non-existent role in reading development *across* skill levels was apparent and that further precluded any theory-based hypothesis or rationale for such an observation) and distinct readings, then such features did not warrant further attention and were eliminated in the larger study.

To accomplish this, redundant observations of a variety of prosodic features were targeted for in-depth analysis using a smaller sample of children of various reading levels. A total of 30 children (33% of the total sample) with complete oral reading data were randomly selected from the larger sample of children participating in the longitudinal study. Selected participants demonstrated a sufficient range of reading ability (25^{th}-99^{th} percentiles) as evidenced by their performance on the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999) taken during the initial spring of first grade assessment. Furthermore, an independent-samples *t* test confirmed that there were no significant differences between the 30 subjects selected for the preliminary analyses and the remainder of the sample (p = .270) on this initial TOWRE measurement. Prosodic measurements were taken from the target GORT passage in the manner described above beginning with the spring of first grade assessment and continuing during the fall and spring of second grade.

The following linguistic features were targeted for spectrographic measurement based on the suggestions of Chafe (1988) and Cooper and Paccia-Cooper (1980) that these structures might require distinct prosodic readings in adults, and the subsequent experimental validation of such readings in adults and children (Clay & Imlach, 1971; Dowhower, 1987; Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004): (1) basic declarative pitch, or fundamental frequency (F_0), declination, (2) intrasentential pause duration, (3) intersentential pause duration, (4) yes-no question pitch rise, and (5) phrase-final comma pause duration. Beyond these previously established measures, additional measurements were made that focused on the intonation contours of key syntactic segments and vowel elongations in sentence-final positions.

The basic declarative sentence-final F_0 declination was determined by isolating the target area on the spectrograph and measuring the pitch change, in Hertz (Hz), from the final pitch peak to the end of the sentence. This was viewed as preferable to simply measuring the fall in pitch on just the final word in the sentence because that measure of declination often fails to describe the fall in pitch heard at the end of a sentence when the final word is unisyllabic (e.g., *A man got out of the car*, where the meaningful difference is noted between the words *the* and *car*). Magnitude of F_0 declination was determined by subtracting the final from the peak fundamental frequency. Measurements were taken on the following three basic declarative example sentences and the mean difference in F_0 was used as an index of sentence-final declination:

1. A man got out of the car.
2. He had a pretty box under his arm.
3. A little girl ran from the house to meet him.

Intrasentential pause duration (in milliseconds) was measured for the presence of pausal intrusions, or inappropriate pauses within words or syntactical units, located in the first three sentences of the text. Pause lengths were determined by visually creating a spectral slice at the limits of the pause interval and noting the duration in milliseconds. Only pause durations exceeding 100 ms were included because they could be reliably measured. In addition, pause measurements were restricted to a maximum duration of

3000 ms as required by general testing protocol established for this study. This restriction was imposed to prevent excessive pausing during oral reading and to facilitate children's passage completion by providing assistance when necessary. Intrasentential pause measurements were taken from the first three sentences of the passage because the sentences were of the same type (declarative) and this kind of analysis is particularly time consuming. Mean pause durations were obtained by averaging across sentences. The actual number of inappropriate, or extraneous, pauses was recorded as well and served as an alternative measurement of intrasentential pausing.

Intersentential pause duration, or the mean length of pauses between sentences, was measured similarly to those within sentences. Intersentential pause lengths were determined by visually demarking the spectrograph at the limits of sentence-final pauses, noting durations (in milliseconds), and averaging across all sentences in the passage. Again, only measurements exceeding 100 ms were included and pause measurements were restricted to a maximum duration of 3000 ms as required by general testing protocol established for this study.

Yes-no question pitch inclination was measured in a similar way to the basic declarative pitch declination; however, where the structure ended with a rise in pitch, F_0 measurements were made from the preceding pitch valley to the final peak. Only one example of this feature type is available in the passage and, consequently, this single measurement served as the index for yes-no question pitch rise.

Phrase-final comma pause duration measurements were determined by slicing the spectrograph at the appropriate phrase boundaries and recording the pause length occurring between the ending of the word preceding the comma and the start of the word following it (e.g., between *Father* and *she*). Measurements were made for each of three phrase-final commas included in the passage and averaged across observations. Duration constraints were identical to those of intrasentential and intersentential pauses.

According to Snow and Coots (1981), intonation contours are regarded "as prominent prosodic markers of the natural units of language" (p. 26). The term *intonation contour* generally refers to the pattern of pitch changes in the voice; however, local intonation contours are those fall-rise patterns that occur specifically at syntactic phrase boundaries within the sentence and at the terminal marker (Dowhower, 1987). Intonation contour was determined by isolating each word in the target sentence and measuring the F_0 at the vocalic nucleus (the voiced portion of the word that produces F_0) of that word. These measurements allowed for the creation of a prosodic profile that provided information about pitch changes over an entire sentence as well as at syntactic boundaries. Measurements were made for each of the first three sentences in the passage. The prosodic profile of each child was correlated with the mean prosodic profile obtained from the adult sample, and the resulting correlation was taken as the child-adult F_0 match for that individual.

Phrase-final lengthening refers to the lengthening of the last stressed syllable of a phrase and is considered a reliable prosodic marker as well (Dowhower, 1987; Klatt, 1975, 1976; Snow & Coots, 1981). Specifically, Cooper and Paccia-Cooper (1980) and Klatt (1975, 1976) demonstrated that fluent speakers and readers elongate stressed vowels most notably in sentence-final positions compared to other phrasal boundaries. Sentence-final vowel elongation was measured by isolating the stressed vowel in the final word of a sentence and recording its voiced duration (in milliseconds). This prosodic feature was measured in each of the first three sentences of the passage and for the yes-no question feature as well. Duration measurements were averaged across sentences and served as an index of sentence-final vowel lengthening.

Once a set of measurements was obtained for all prosodic features across the three time points, we were able to examine the target readings and growth trends of each individual structure and minimize the total number of prosodic features (based on the criteria explained above) that would be targeted for more thorough study. To analyze this preliminary data, a person-period data set was created which contained measurements of each prosodic feature for each of the three waves of data collection for all 30 subjects. In this exploratory analysis we were interested in (1) determining which prosodic structures demonstrate sensible patterns of change over time, and (2) examining the particular change trajectories of each feature to both characterize the specific shape of the developmental trend and to make certain that the trends were consistent with established target readings. This can be accomplished in a number of different ways. Careful examination of the sample descriptive statistics was carried out for each structure to ensure that the values obtained across time points were of expected directionality (e.g., pitch inclination vs. declination, short duration vs. long duration) and that the measurements were indicative of an appropriate change trajectory. Each prosodic feature was separately examined for trend by means of repeated measures ANOVA as well. We constructed orthogonal polynomial contrasts that were useful in determining the degree of change over time and the relative contribution of each polynomial component (i.e., linear, quadratic, cubic) of the trend. In addition, we also carried out consecutive time comparisons, or “profile” contrasts, to determine whether each consecutive time point was significantly different from the immediately previous time point. Brief descriptive accounts were presented for each feature (where appropriate) as well. Our reasons for doing so were based on the beliefs that (1) a more complete understanding of developmental prosody (at least as it concerns the variables currently under investigation) could be obtained through qualitative examination of emergent patterns in the data, and (2) procedural observations (e.g., those concerning appropriate measurement techniques) obtained throughout the initial measurement process could also be necessary for making informed selections decisions regarding the specific prosodic features to be included in the full study. Taken together, the combination of sample statistic and repeated measures analyses, along with qualitative descriptions affords a comprehensive presentation of the preliminary data useful for the purpose of variable selection. In order for a feature to be eliminated from further analysis, either a lack of change or a theoretically unjustified change would need to be evident across subjects, and/or a particular prosodic interpretation would need to be unclear across subjects or inconsistent with the established target. Furthermore, feature elimination due to issues of measurement would be considered a justification as well.

Selection Decisions Based On the Findings from the Preliminary Analysis of Prosodic Features

Means, standard deviations, and ranges for each prosodic variable at each time point can be found in Table 2.1. Plots for each prosodic feature can be seen in Appendix A. While the majority of features displayed mean values that were of appropriate magnitude, directionality, and indicative of change over time, a review of both the descriptive statistics and results from the repeated measures analyses suggested some need for variable modification and/or elimination.

Mean measurements for the sentence-final F_0 declination were appropriately expressed as negative values and displayed a general pattern of change in which the magnitude of the declination increased over time. Such observations are theoretically sound and consistent with previous research suggesting that skilled readers end declarative sentences with marked declinations in pitch, with the size of the declination varying as a function of skill (Clay & Imlach, 1971; Dowhower, 1987; Miller & Schwanenflugel, 2006; Schwanenflugel et al, 2004). Examination of the results of the repeated measures analysis revealed a significant linear developmental trend during grades 1 through 3, F (1, 29) = 6.087, p = .020, partial eta^2 = .173. The effect size, however, was rather small. Additional time comparisons revealed that the mean difference between performances at times 1 and 2 was not significant (-1.578, p = .744) while the mean differences in performance between times 1 and 3 (11.422, p = .020) and times 2 and 3 (13.000, p = .020) were both significantly different from each other. Thus, while the development of sentence-final pitch declination appears to proceed in a generally linear manner with the magnitude of the declination becoming larger as children become more skilled, performance during the fall of second grade (time 2) represents something of a disruption in progress as children in our sample may have experienced a minor loss in this type of prosodic interpretation between school years. One particular observation regarding the sentence-final declination, however, was rather interesting. Although this feature demonstrated an expected pattern of growth, the magnitude of the initial measurement (-25 Hz) was much larger than we would have anticipated. Therefore, it may be that sizeable declinations in pitch at the ends of declarative sentences are in fact evident in the oral reading of children from an early age and increase somewhat over time, meaning that the developmental trend may not necessarily originate from a relatively flat, or non-existent, prosodic interpretation.

Investigation of intrasentential pausing consisted of two separate (though related) measurements: a calculation of the average pause duration in milliseconds and a tally of the total number of pausal intrusions made during oral reading. Our reasons for doing this were due in large part to a general uncertainty as to which method might provide the best measurement and a concern over the potential for reporting misleading results when using the pause duration measurement as opposed to the total number of pauses. A brief example is included for clarification purposes. Suppose Child A's reading consisted of 7 extraneous pauses with durations of 172, 312, 236, 1293, 482, 647, and 125 milliseconds, the average of which would be 467 ms. Alternatively, Child B recorded only 1 extraneous pause of 500 ms in duration. Given this information few could argue that the two readers are more or less equivalent in terms of fluency (a child whose reading is characterized by 7 interruptions would certainly be considered less fluent than a child who read with only 1 interruption); however, it would appear that the two readers performed similarly when the mean duration value is presented alone. Furthermore, a review of the descriptive statistics showed negligible declinations in intrasentential pause duration over time (a mean difference of 95 ms between times 1 and 3). The total number of pausal intrusions, on the other hand, appeared to decline steadily over the course of development. In fact, results of the repeated measures analysis for this particular feature demonstrated a significant linear trend and a moderate effect size, F (1, 29) = 29.523, $p < .001$, partial eta^2 = .504. Furthermore, pairwise comparisons revealed significant mean differences between all time point combinations (mean difference = 1.400, p = .007 for time 1-time 2; mean difference = 3.433, p <.001 for time 1-time 3; mean difference = 2.033, p = .003 for time 2-time 3). That a reduction in the number of pausal intrusions is evident over the course of development is somewhat expected; however, the strength and clarity of this finding serves as an important theoretical illustration and highlights a potential connection between skill development (i.e., the development of automatic decoding and word reading skills) and prosody. Consequently, in light of the information presented above the total number of pauses was retained as an indicator of intrasentential pausing and the duration measurement was discarded.

An examination of the descriptive statistics for both the intersentential and phrase-final comma pause durations revealed general patterns of change over time in which the mean pause values decreased throughout the course of development. The observed trends for these features were indicative of prosodic renderings that became increasingly consistent with the appropriate "target" readings as determined in previous research (see Dowhower, 1987; Miller & Schwanenflugel, 2006; Schwanenflugel et. al., 2004). Intersentential pause duration measurements displayed a linear trend over time and a small to moderate effect size as well, F (1, 32) = 11.039, p = .002, partial eta^2 = .256. Although children's between-sentence pause durations decrease consistently from the initial measurement taken during the spring of 1st grade, only the extreme difference between the means at time 1 and time 3 was significant (147.47, p = .002). Mean differences between times 1 and 2 and times 2 and 3 were not significant (47.57, p = .547 and 100.36 p = .173, respectively). Results for the phrase-final comma pause feature were similar to those of the other pause features, demonstrating a significant linear trend and a small to moderate effect size, F (1, 32) = 12.199, p = .001, partial eta^2 = .276. Review of the time comparisons revealed that while the mean difference in performance between times 1 and 2 was not significant (43.36, p = .406), differences between times 1 and 3 (189.88, p = .001) and times 2 and 3 (146.52, p = .006) were significant. Overall, this finding suggests that for young readers commas may be viewed as obligatory signals to pause; however, the observed decease in pause duration over time could indicate that children may no longer feel driven to mark every comma with a pause as their reading skills develop. Having met the criteria for inclusion in the full study (i.e., demonstration of theoretically sound patterns of change and evidence of distinct prosodic readings), the intersentential and phrase-final comma pause durations were included for subsequent analyses.

Children's readings of the yes-no question feature were somewhat similar to those of the sentence-final declination in that the appropriate prosodic interpretation (a sizeable pitch inclination) was present at the initial measurement. However, valid concerns regarding the acceptability and reliability of this feature relative to the criteria established for variable selection were present. Specifically, although the obtained values were of appropriate directionality across measurements (positive values reflecting a pitch inclination), the developmental profile was relatively erratic with the largest magnitude recorded during the spring of 1st grade assessment. Performance appeared to dip considerably during the fall of the 2nd grade school year (perhaps reflecting a loss of ability over the summer between school years) and failed to fully recover by the end of the 2nd grade. Not only did the results of the repeated measures analysis fail to yield a significant developmental trend *F* (1, 29) = .473, $p = .497$, partial eta^2 = .016, but the pairwise comparisons revealed that there were no significant differences between any of the time point combinations as well (mean difference = 9.800, $p = .201$ for time 1-time 2; mean difference = 4.967, $p < .497$ for time 1-time 3; mean difference = -4.833, $p = .344$ for time 2-time 3). Furthermore, analysis of the yes-no question was considered problematic from a measurement perspective as well, given that the selected GORT passage contains only one observation of this feature. Measurements for all other prosodic variables are obtained by averaging over redundant observations, a procedure which we feel adds to the reliability of the measurement. Taken together, the absence of consistent developmental change and the potential for inconsistencies in measurement were grounds for eliminating the yes-no question features from further consideration.

Examination of phrase-final vowel elongation yielded results similar to the yes-no question feature in that this variable failed to demonstrate a pattern of change over time and was characterized by significant measurement difficulties as well. Specifically, the growth trajectory was flat, or non-existent, as duration measurements across time points were found to be within 10 ms of each other. Repeated measures analysis confirmed the absence of a significant growth trend *F* (1, 29) = 2 .573, $p = .120$, partial eta^2 = .082, and indicated that only the mean difference between Time 2 and Time 3 was significant (-8.000, $p = .020$; the mean differences between Time 1-Time 2 and Time 1-Time 3 were not significant with obtained *p* values of .332 and .120, respectively). Furthermore, we experienced considerable difficulty isolating the vowel sounds and measuring their elongations given that the durations were rather minute (often less than 100ms) and the vowels frequently blended with surrounding sounds. Our lack of experience with this type of measurement in the context of oral reading meant that we had no clear indication of what the “target” reading should be or what constituted an acceptable measurement value as well. Due to the noted lack of developmental change and concerns about our ability to obtain reliable measurements for this feature it was necessary to remove phrase-final vowel elongation from the study.

Finally, analysis of the mean values for the child-adult F_0 contour match across measurement points revealed a clear pattern of change in which the overall prosodic envelope of children’s oral reading increasingly approximated that of adults with the passage of time. The initial correlation between child and adult prosodic contours was moderate (.527); however, the correlations increased steadily throughout the course of development suggesting that there is in fact a general “target” prosodic contour to which developing readers strive. Repeated measures analysis of the child-adult F_0 match demonstrated a significant linear trend over time, F (1, 29) = 26.149, $p < .001$, partial eta^2 = .474, and, in the context of the other prosodic features analyzed, a somewhat larger effect size was observed for this feature as well. Additional time comparisons revealed significant mean differences between times 1 and 3 (-.162, p <.001) and times 2 and 3 (-.093, p = .016); however, the mean difference between performances at time 1 and at time 2 was not significant (-.069, p = .136). Overall, we were satisfied that this variable met the criteria for retention.

In sum, this preliminary phase of the study served an important function in that it allowed us to identify those prosodic features that were relevant for a developmental study of prosodic reading. The sentence-final F_0 declination, intrasentential pause total, intersentential pause duration, phrase-final comma pause duration, and child-adult F_0

contour match demonstrated significant linear patterns of change over time that were both distinct and consistent with the "target" prosodic interpretations established in previous research (see Clay & Imlach, 1971; Dowhower, 1987; Miller & Schwanenflugel, 2006; Schwanenflugel et. al., 2004). Consequently, they were retained for further analysis. Furthermore, this characterization of the development of prosodic text reading is of particular theoretical importance in that it suggests that prosody is a continuously evolving feature of oral reading with an obvious presence early on in the reading process. Absence of evidence showing smooth and consistent change over time combined with significant measurement issues resulted in removal of the yes-no question and phrase-final vowel elongation features.

Procedures for Longitudinal Analysis of the Development of Reading Prosody and 3rd Grade Outcomes

In this second phase, attention was restricted to those features identified in the preliminary analysis as showing evidence of developmental change and distinct prosodic readings (sentence-final F_0 declination, number of intrasentential pauses, intersentential pause duration, phrase-final comma pause duration, and child-adult F_0 contour match). The reading prosody of all 92 children participating in the longitudinal study was then examined beginning with the spring of first grade assessment and continuing through the fall and spring of second grade time points. An additional outcome assessment, conducted during the spring of the third grade school year, was included in this large-scale study as well. Outcome measures included formal assessments of oral reading fluency and reading comprehension. The design for this analysis can be seen in Figure 2.3.

The purpose here was to characterize the onset and development of prosodic reading throughout the years of primary reading acquisition (grades 1 through 3) in relation to the acquisition of various component reading skills (e.g., word reading efficiency, oral reading fluency, reading comprehension), and with explicit comparisons to theoretical accounts of reading development, in general, and prosodic development, in specific. Longitudinal data affords not only the opportunity to observe how the development of reading prosody proceeds, but also the opportunity to confirm or create hypotheses that may be of substantial theoretical and practical importance. For example, skill-based differences in the expressiveness aspect of same-age children's reading support the possibility that prosody may develop after successful accumulation of various lower-level skills. Furthermore, additional outcome measures would demonstrate the general effect of prosodic reading ability on performance and permit comparisons between the development of reading prosody and other aspects of reading achievement.

Outcome Reading Assessments. We added one additional time point to the design of our study which was comprised of the assessments administered during the spring of third grade time point from the larger longitudinal study of oral reading fluency. The reason for including these outcome measures is based upon our stated objectives to examine how the early development of reading prosody predicts both the ability to read fluently and reading comprehension skills at the end of grade 3. For these outcome assessments the following measures were given:

(a) *Oral reading fluency outcome assessment.* The Gray Oral Reading Tests, Fourth Edition (GORT-4)--Form B was administered to obtain an estimate of reading fluency skill in reading connected text. Form B does not include the passage targeted for prosodic analysis.

(b) Reading comprehension outcome assessment. Children were administered the Reading Comprehension subtest of the Wechsler Individual Achievement Test (WIAT) to obtain an independent estimate of reading comprehension skill.

Table 2.1

Descriptive Statistics for Preliminary Analysis of Prosodic Features

Prosody Feature	*N*	Minimum	Maximum	*M*	*SD*
Basic Declarative F_0 Change (Hz)					
Time 1	30	-66	-1	-25	16
Time 2	30	-69	7	-23	20
Time 3	30	-82	-15	-36	21
Intrasentential Pause Duration (ms)*					
Time 1	30	0(0)	1419(21)	404(5)	267(5)
Time 2	30	0(0)	1640(20)	383(4)	348(4)
Time 3	30	0(0)	1110(9)	309(2)	279(2)
Intersentential Pause Duration (ms)					
Time 1	30	311	1765	637	331
Time 2	30	113	1672	575	440
Time 3	30	78	1388	479	297
Yes-No Question F_0 Change (Hz)					
Time 1	30	-13	189	29	41
Time 2	30	0	90	19	18
Time 3	30	0	111	24	25
Phrase-final Vowel Duration (ms)					
Time 1	30	76	115	88	9
Time 2	30	61	135	85	17
Time 3	30	74	112	93	10
Phrase-final Comma Pause Duration (ms)					
Time 1	30	0	1109	314	296
Time 2	30	0	1475	300	304
Time 3	30	0	684	140	154
Child-Adult F_0 Contour Match					
Time 1	30	.041	.757	.527	.203
Time 2	30	.170	.864	.595	.166
Time 3	30	.463	.836	.688	.094

Note. Values included in parentheses reflect the descriptive statistics for the raw number of intrasentential, or extraneous, pauses recorded during oral reading. Time 1 = Spring of 1st grade, Time 2 = Fall of 2nd grade, Time 3 = Spring of 2nd grade as outlined in the assessment design for the measurement of prosodic text reading.

Figure 2.1

Layout and assessment design for the developmental study of oral reading fluency

Time 1	Time 2	Time 3	Time 4	Time 5
Spring 1st Grade	Fall 2nd Grade	Winter 2nd Grade	Spring 2nd Grade	Spring 3rd Grade
TOWRE-Form A GORT-3, Form A WIAT-RC	TOWRE-Form A GORT-4, Form A WIAT-RC	TOWRE-Form B GORT-4, Form B WIAT-RC	TOWRE-Form A GORT-4, Form A WIAT-RC	TOWRE-Form B GORT-4, Form B WIAT-RC

Figure 2.2

Selected assessment points for the measurement of prosodic text reading

Time 1	Time 2	Time 3
Spring 1st Grade	Fall 2nd Grade	Spring 2nd Grade
TOWRE-Form A GORT-3, Form A WIAT-RC	TOWRE-Form A GORT-4, Form A WIAT-RC	TOWRE-Form A GORT-4, Form A WIAT-RC

Note. Passage 1 on GORT-3, Form A and Passage 3 on GORT-4, Form A are identical.

Figure 2.3

Layout and assessment design for the full developmental study of prosodic text reading (Phase 2)

Time 1	Time 2	Time 3	Time 4
Spring 1st Grade	Fall 2nd Grade	Spring 2nd Grade	Spring 3rd Grade Outcome Assessment
Prosodic measurement on GORT-3, Form A (passage 1)	Prosodic measurement on GORT-4, Form A (passage 3)	Prosodic measurement on GORT-4, Form A (passage 3)	Sum of passage fluency scores on GORT-4, Form B
TOWRE-Form A WIAT-RC	TOWRE-Form A WIAT-RC	TOWRE-Form A WIAT-RC	WIAT-RC

Note. Passage 1 on GORT-3, Form A and Passage 3 on GORT-4, Form A are identical.

CHAPTER 3

ANALYSES AND RESULTS

Analyses were specifically designed for the purpose of achieving the following four research objectives: (1) to determine whether decreases in pause duration over time serve a causal function for the development of larger pitch changes; (2) to determine the extent to which the growth of prosody during grades 1 and 2 is predictive of oral reading fluency in grade 3; (3) to determine whether the growth of reading prosody during grades 1 and 2 is predictive of comprehension skill in grade 3; and (4) to determine the extent to which the development of reading prosody adds to our ability to account for reading fluency and comprehension outcomes beyond word reading speed and accuracy.

Prior to conducting any analyses we created scatter plots and screened for outliers on all prosodic variables. Kline (2005) presents a common "rule of thumb" for identifying univariate outliers in which scores that are more than 3 standard deviations from the mean are considered extreme. Using this standard we determined that 12 subjects displayed an extreme value for one of the five prosodic variables in comparison to the mean value for that variable at a given measurement point. Extreme values were dispersed across time points and variables as follows: 1 subject on the sentence-final declination at Time 1, 2 subjects on the internal comma pause at Time 1, 2 subjects on the intersentential pause at Time 2, 2 subjects on the internal comma pause at Time 2, 1 subject on the F_0 match at Time 2, 1 subject on the sentence-final declination at Time 3, 1 subject on the intersentential pause at Time 3, 1 subject on the internal comma pause at

Time 3, and 1 subject on the F_0 match at Time 3. Since statistics derived from data that include outliers will often be misleading, each of these values was constrained to 3 standard deviations from the mean. Means, standard deviations, and ranges for each prosodic variable at each time point for the complete longitudinal sample can be found in Table 3.1.

Given the absence of a single systematic developmental study of reading prosody in the research literature, the first objective, aimed to place the development of prosodic text reading in the context of the development of related reading skills. Chall (1996b) and Kuhn and Stahl (2003) suggest that children must first solidify decoding and word reading capabilities prior to rendering the rhythmic and tonal aspects of language during oral reading. Thus, decreases in pause duration may in fact reflect successful acquisition of word reading skills rather than serve as a true indicator of reading prosody. Our goal was to examine whether pause and pitch features emerge simultaneously or whether pause features are first to develop and play a causal role in the subsequent development of pitch changes.

The second and third objectives concerned the extent to which the cumulative effects of the development of reading prosody predicted fluent reading and comprehension skills in grade 3, a year later. Specifically, we were interested in determining what separate impact was evident for pause and pitch on these outcomes. Presumably, if prosody is evidenced primarily by changes in pitch, rather than by decreases in pause duration, then we might expect pitch to be a stronger predictor of both fluency and reading comprehension (given that prosody is considered a higher-order component of fluency and that fluency is so closely related to comprehension skill).

However, it is also possible that the impact of the development of pause and pitch variables might vary depending on the particular outcome under consideration.

Finally, for the fourth objective, we were interested in examining the relative contributions of the development of pause and pitch features on both fluency and comprehension outcomes beyond the impact of word reading speed and accuracy. Essentially, our goal was to determine the extent to which the development of reading prosody adds to our ability to account for later fluency and comprehension skills when word reading skills are considered as well. If word reading skill eliminates or alters any of the previously established relationships between the development of prosody and various reading outcomes (objectives 3 and 4) then we may conclude that prosody has relatively little, or only very specific, additional impact on fluency and comprehension. If, however, it becomes apparent that either the development of pause or pitch, or both, continues to exert a significant influence on reading outcome measures despite the presence of word reading skills, then we may conclude that the development of reading prosody is a true additional predictor of fluency and comprehension skill in grade 3.

Analyses were conducted in sequential steps related to the four stated objectives and where each set of new analyses built upon those that were carried out previously. Path modeling procedures (a variant of the SEM approach used to provide estimates of the magnitude and significance of hypothesized causal relationships between two or more variables) were identified as the methods through which the achievement of these objectives would be possible. However, before carrying-out these path model analyses a more thorough review of our data was required.

Variable Selection, Description of Data, and Criteria Specification for Path Model Tests

Given that the validity of the SEM approach depends on meeting certain assumptions regarding the data that are analyzed, a presentation of basic descriptive information about the data on which an SEM analysis is based is essential (Hoyle & Panter, 1995). Specifically, such a description should include information about the distributions of individual variables and the multivariate distribution of the variables in the model(s) to be estimated. Furthermore, Tanaka (1993) and Hoyle and Panter (1995) suggest that authors specify which indexes of overall fit are to be presented and include an interpretation of said indexes prior to reporting the results of SEM analyses as well.

Variable selection. Sample size considerations mandated that we simplify our approach to modeling prosodic development considerably for the remaining analyses. According to Kline (2005), with fewer than 100 cases almost any type of SEM analysis could be considered untenable unless the model evaluated is rather simplistic in its design (pg. 15). We simplified our design in two ways:

First, we eliminated our second prosodic measurement (taken during the fall of 2nd grade) and created an adjusted design with evenly spaced measurement intervals of approximately 1 year in duration. The adjusted model focused on prosodic measurements taken during the spring of 1st grade and spring of 2nd grade assessments with an additional outcome assessment completed during the spring of the 3rd grade school year. The adapted design can be seen in Figure 3.1. Furthermore, along with changes in design structure, it was necessary to focus on observed, rather than latent, prosodic variables and reduce the overall number of variables used in these analyses as well. In pursuit of model simplicity we elected to use single indicators of pause and pitch. Consequently, we selected the intrasentential pause and child-adult F_0 contour match variables for all remaining model tests. These particular prosodic features were selected because ANOVAs had suggested that they were the most robust features included in this developmental study of prosodic text reading with significant linear trends at the $p < .001$ level and effect sizes of .504 and .474, respectively. In addition, preliminary factor analyses of all prosodic variables using a principle components extraction and Varimax rotation demonstrated that these two variables were high loaders (above .800) on their respective factors. Significant correlations were found between measurements at the spring of 1st grade and spring of 2nd grade time points for each of these features as well (.773 between intrasentential pause measurements and .363 between F_0 match measurements). Furthermore, on conceptual grounds, intrasentential pause and F_0 match are fairly "global" indicators of pause and pitch. Specifically, the F_0 match represents the overall prosodic profile and incorporates sentence-final declination within its measurement (results of Schwanenflugel et al. demonstrated F_0 match to be a significant predictor of comprehension skill as well) while the intrasentential pause accounts for all irrelevant pauses unconnected to punctuation. Finally, both the F_0 match and the intrasentential pause measurements were collected from the same segment of text in the selected GORT passage.

Description of data. LISREL 8.8 for Windows (Joreskog & Sorbom, 2006) was used for preliminary data screening procedures and to conduct all remaining model tests. Means, standard deviations, and skewness and kurtosis values for all variables included in the model tests are presented in Table 3.2. In addition, correlations between all variables are presented in Table 3.3. A review of the univariate summary statistics for continuous variables revealed that the means and standard deviations were of generally expected values. Furthermore, analysis of the skewness and kurtosis values indicated that all variables were approximately normally distributed (values less than |2|).

Normality of the individual variables, however, does not guarantee multivariate normality. The joint distributions of all combinations of variables must be normal as well (a condition that is assumed for most estimation procedures used in SEM). Since nonnormality can result form the presence of outliers in the data, DeCarlo's (1997) macro was used to screen for multivariate outliers prior to screening for multivariate normality. According to Kline (2005), it is possible for a case to be considered a multivariate outlier if the pattern among variables for a given subject is atypical in the sample (despite the absence of extreme values on any variable individually). Based on the 5 observations with the largest Mahalanobis distances, a measure of how far an observation's values on the variables are from the multivariate mean of all variables, it appears that there were no single multivariate outliers in the data set (all values below the critical *F* of 16.34). Finally, to assess the assumption of multivariate normality PRELIS outputs a value called the "relative multivariate kurtosis" statistic. General recommendations suggest that values < 2.0 are indicative of normal multivariate distributions and, consequently, the obtained relative multivariate kurtosis value of 1.26 indicated approximate multivariate normality with the current data.

Treatment of missing data (approximately 18% in the entire data set) was addressed through the method of multiple imputations using the Expectation-Maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977), a general iterative algorithm for computing maximum likelihood estimates from incomplete data. The EM algorithm is used to obtain estimates of population means and covariances which LISREL then uses to obtain starting values for the maximum likelihood procedure. Although ML is the default method of estimation in LISREL, it is the most widely researched estimator and it works well under a variety of conditions (e.g., small sample size) as well. Furthermore, characteristics of the data, such as approximate multivariate normality (described previously), suggest that ML is the appropriate estimation procedure.

Fit Criteria. Based on the combination of recommendations proposed by Hoyle and Panter (1995) and Hu and Bentler (1999) the following fit indexes were chosen to evaluate model fit: Minimum Fit Function Chi-square, Goodness of Fit Index (GFI), Non-Normed Fit Index (NNFI), Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA). The LISREL GFI is a measure of the proportion of the observed covariation that is accounted for by the model. Hoyle and Panter (1995) recommend including this index to complement presentation of the χ^2 statistic as measures of overall fit. A GFI value exceeding .95 is considered an indication of model fit. In addition to indexes of overall fit, Hoyle and Panter (1995) recommended inclusion of incremental fit indexes (specifically Type 2 and Type 3 indexes) in order to assess model fit with reference to some type of baseline model (something that the chi-square value does not indicate). Type 2 indexes represent the proportion of increased fit the hypothesized model shows over the null model and incorporate the expected values of the chi-square under the central chi-square distribution. The NNFI was selected as a Type 2 index using a cut-off value of at least .95. Type 3 indexes represent the proportion of increased fit the hypothesized model shows over the null model and incorporate the expected values of the chi-square under the noncentral chi-square distribution. The CFI, selected as a Type 3 index, compares the noncentrality parameters of the target and baseline models. The ML-based estimate of the CFI is especially preferred when sample size is small as well. Again, Hu and Bentler (1999) recommend a CFI value of .95 or above. Finally, the RMSEA is a highly recommended fit index that provides a standardized measure of the lack of fit of the population data to the model. Desired values of this index are low and Hu and Bentler (1999) recommend a cut-off close to .06.

With a thorough description of our data complete (demonstrating that we have met the underlying assumptions of SEM concerning the distributional characteristics of the data) and criteria for evaluating model fit established, we proceeded with our remaining objectives and model tests.

Analysis of Prosodic Development in Sequence and in the Context of Reading Development

Our first objective was to examine the relationship between pause and pitch features over time using several logically-based theoretical models. Our primary interest was in determining whether decreases in the presence of pausal intrusions over time were a necessary initial condition for the onset and development of more adult-like pitch contours to occur, or whether these prosodic features grow more or less in unison. In this analysis, we focused on those prosodic features (intrasentential pause and child-adult F_0 contour match) that were considered the most robust indicators of pause and pitch as noted in the variable selection section. Path diagrams of all theoretical models to be tested can be found in Figure 3.2.

The first model tested, called the Independence model, was essentially a null model which assumed that no cross-lagged effects would be found between the number of pausal intrusions present during oral reading and the development of pitch contour at subsequent time points, and instead implies that pausing will only contribute to pausing at the subsequent time point and pitch contour will only contribute to pitch contour at the subsequent time point. The Independence model was examined first because it was the null model and represented the benchmark used to evaluate ensuing model tests. As expected, each path included in this model was significant at the $p < .01$ level. This finding suggested that the number of pausal intrusions made during the spring of 1st grade was significantly related to the number of pausal intrusions observed during the spring of the 2nd grade school year. Similarly, the child-adult F_0 contour match measured during the spring of 1st grade was significantly related to the F_0 contour during the spring of 2nd grade. Overall model fit to the observed data, however, was considered marginal, χ^2 (3) = 8.25 (p = .041), GFI = 0.96, NNFI = 0.90, CFI = 0.95, RMSEA = 0.14. Path weights, standard errors, and *t* values for the Independence model are summarized in Table 3.4.

According to Chall (1996b), as reading becomes increasingly less halting, perhaps as evidenced by decreases in the number of pausal intrusions that result from the acquisition of automatic decoding skills, children develop the ability to represent what is read in conversational tones. This would suggest that automatic decoding, and, consequently, decreases in pausing, might precede the development of pitch changes and perhaps serve a causative function for the emergence of pitch features as well. In our second model, labeled the Skill Development model, we hypothesized that pausing may be more related to difficulties with decoding at the word level, and that the presence of a substantial number of pauses during text reading reflects a level of dysfluency that precludes the development of true prosodic interpretation as evidenced by fluctuations in pitch. Thus, once pausing is of minimal occurrence, thereby indicating the achievement of speeded and accurate oral reading, we should observe the development of more appropriate pitch contours during oral reading.

Results supported this skill development hypothesis and the overall fit of the model to the observed data was excellent, χ^2 (2) = 1.27 (p = .53), GFI = 0.99, NNFI = 1.02, CFI = 1.00, RMSEA = 0.0. Whereas the initial measurements for pausal intrusions and F_0 contour continued to demonstrate significant influences on their respective measurements at the subsequent time point ($p < .01$), the additional path between pause at time 1 and pitch contour at time 2 was significant at the $p < .01$ level as well. Essentially, this finding suggests that there is an inverse relationship between the number of pausal intrusions observed during the spring of 1st grade measurement and children's ability to produce adult-like pitch contours during the spring of the 2nd grade school year. A reduction in the number of inappropriate pauses, a reflection of the achievement of speeded and accurate oral reading, appears to impact the subsequent measurement of prosody in pitch. A summary of the path weights, standard errors, and t values for the Skill Development model is presented in Table 3.4. The chi-square difference test (χ^2_{diff}) was then applied for comparison of fit between the Independence model (the most restrictive model representing the null hypothesis) and the Skill Development model (a less restrictive model). The resulting statistic illustrated a significant difference between the models and a rejection of the null hypothesis, χ^2_{diff} (1) = 6.98, $p < .01$. Thus, the inclusion of the additional skill development path resulted in a significant improvement over the Independence model given the difference in degrees of freedom.

We also considered a third model, labeled the Reciprocal Effect model, which suggested that a cross-lagged relationship may exist between the number of pausal intrusions and the child-adult F_0 match correlation. A cross-lagged relationship would be evident in the current investigation if pausing at the initial time point impacted pitch contour at the subsequent time point *and* pitch contour had a significant impact on pausing at the subsequent time point as well. In other words, the two variables might have a reciprocal effect, where appropriate pitch profiles are dependent upon decreases in the number of pausal intrusions, and decreases in the total number of pausal intrusions observed during oral reading are dependent upon previous pitch profiles. Results demonstrated that the paths of the nested Skill Development model were significant at the $p < .01$ level; however, the additional path between F_0 contour at time 1 and pause at time 2 (representing the reciprocal effect) was not significant. A summary of the path weights, standard errors, and *t* values for the Reciprocal Effect model are summarized in Table 3.4. Although a review of the fit indices suggests exceptional model fit to the observed data, χ^2 (1) = 0.18 (p = .67), GFI = 1.00, NNFI = 1.04, CFI = 1.00, RMSEA = 0.0, the lack of significance for the reciprocal path between the initial pitch contour measurement and the number of pausal intrusions at the subsequent time point indicates a lack of support for the reciprocal effect theory. Furthermore, comparison between the Reciprocal Effect and Independence models revealed a significant difference $\chi^2_{diff}(2) = 8.07$, $p < .05$, indicating better fit than the null model, given the cost in degrees of freedom.

Comparison between the Reciprocal Effect and Skill Development models, however, revealed no significant difference between models, $\chi^2_{diff}(1) = 1.09$, $p = .30$.

Taken together, the model tests and comparisons demonstrated that both the Skill Development and Reciprocal Effect models were better representations of the observed data than the Independence, or null, model. However, the Skill Development model emerged as the best fit to our observed data given that the lack of significance noted for the reciprocal path between pitch contour at time 1 and pause at time 2 failed to support the reciprocal effect hypothesis.

Analysis of Prosody as a Predictor of Reading Fluency and Reading Comprehension

The model tests for objectives 2 and 3 addressed the extent to which the cumulative effects of pause and pitch development were subsequently related to outcome measures of oral reading fluency and reading comprehension, respectively. Each model used the Skill Development model as a starting point to explain the relationship between the pause and pitch aspects of prosody, and then examined the individual contribution of the pause and pitch variables as predictors of later achievement. Given that prosodic reading is widely considered one of the defining features of fluent reading we would expect the development of prosody to be predictive of fluency at the end of third grade. However, what impact the rate of prosodic development may have on fluency achievement is unknown. If decreases in pausing precede the development of pitch changes as Chall (1996b) and Kuhn and Stahl (2003) have suggested (and as was demonstrated in previous model tests), it may be that the cumulative effects of the early development of pause structures predict fluency in a different way than later changes in pitch. Moreover, should decreases in pausing serve simply as a proxy for the achievement of automatic word decoding skills, leaving pitch changes as the true indication of prosodic text reading, then it could be that changes in pitch might be more predictive of fluent reading than are decreases in pausing. In terms of reading comprehension as an outcome, Miller and Schwanenflugel (2006) demonstrated that although skill based differences were evident in both the pause and pitch aspects of prosody, pitch fluctuation was related to increased reading comprehension skill in third grade children whereas pause duration was not. Yet, it remains to be seen whether the rate of prosodic development has any impact on later comprehension skills. If in fact the development of changes in pitch appears to be indicative of greater levels of fluency, it follows that we would also find pitch changes to be more predictive of greater levels of comprehension as well (given the strong relationship between fluency and comprehension). The path diagrams for these models can be found in Figures 3.3 and 3.4.

As expected, paths of the nested Skill Development model within the Fluency as Predicted Outcome model were significant at the $p < .01$ level. Furthermore, results demonstrated that while the indirect cumulative effect of decreases in the number of pausal intrusions over time was a significant predictor of fluency in grade 3 ($p < .01$) the indirect cumulative effect of the development of appropriate pitch contour was not significantly related to fluency as an outcome measure. Given that definitions and measures of oral reading fluency often consist of reading rate and accuracy components it is not surprising that a decrease in the number of pausal intrusions (perhaps a reflection of the achievement of automatic decoding and word reading skills) predicts later fluency. However, in light of previous research by Miller and Schwanenflugel (2006) and Schwanenflugel et al. (2004) which demonstrated that pitch features were related to increased reading comprehension skills (whereas pauses were not) it is somewhat surprising that the development of pitch contour is not predictive of fluency as well (considering the strong relationship between fluency and comprehension). There are, however, critical differences between the current investigation and those aforementioned studies particularly concerning the developmental nature of the present examination of prosody and the fact that previous studies used somewhat older groups of children with prosodic and outcome measurements that were taken concurrently. Moreover, since previous analyses have suggested that decreases in pausing appear to precede the development of pitch, it may be that pitch contour could become a significant predictor of fluency later on in development. Additional results of the model test, however, indicated that the direct relationships between both pitch contour and pause at time 1 and fluency at time 3 were significant at the $p < .01$ level. Thus, it is also possible that in terms of pitch contour the developmental aspect may not be as important as initial ability. The overall fit of the Fluency as Predicted Outcome model to the observed data was excellent, χ^2 (2) = 1.27 ($p = .53$), GFI = 0.99, NNFI = 1.02, CFI = 1.00, RMSEA = 0.0. Path weights, standard errors, and *t* values for the Fluency as Predicted Outcome model are summarized in Table 3.5

The Comprehension as Predicted Outcome model shared many similarities with the previous model discussed. Again, the skill development paths were significant at the $p < .01$ level. Furthermore, whereas the indirect cumulative effect of decreases in pausal intrusions over time was a significant predictor of comprehension skill in grade 3 ($p < .01$), the indirect cumulative effect of the development of appropriate pitch contour was not a significant predictor of later comprehension skill. However, the Comprehension as

Predicted Outcome model differed from the Fluency as Predicted Outcome model in that while the direct relationship between F_0 contour match at time 1 and comprehension at time 3 was significant ($p < .01$), the direct relationship between the number of pausal intrusions at time 1 and comprehension at time 3 was not significant. Thus, it appears that while both the pause and pitch variables predict comprehension skill in grade 3, the manner in which they do so differs considerably. The absence of a significant cumulative effect of pitch development on comprehension suggests either that initial pitch contour (and not the rate of development) is the most crucial indicator of later comprehension skill and that good comprehenders use pitch as a comprehension mechanism from the outset, or that the hypothesized delay in the development of pitch relative to the development of pause structures (skill development hypothesis) precludes a significant finding given the time frame of this study. Alternatively, cumulative decreases in the number of pausal intrusions during oral reading (a developmental trend indicating possible acquisition of automatic decoding and word reading skills) predicts later comprehension skill, whereas the initial number of pausal intrusions does not. The overall fit of the Comprehension as Predicted Outcome model to the observed data was exceptional and identical to that of the Fluency as Predicted Outcome, χ^2 (2) = 1.27 (p = .53), GFI = 0.99, NNFI = 1.02, CFI = 1.00, RMSEA = 0.0. A summary of the path weights, standard errors, and *t* values for the Comprehension as Predicted Outcome model can be found in Table 3.6.

Analysis of the Relationship between Prosody, Word Reading Efficiency, and Reading Fluency and Reading Comprehension Outcomes

The remaining model tests, for objective 4, concerned an examination of the individual contributions of both the development of prosodic text reading and word reading abilities in predicting oral reading fluency and reading comprehension during the spring of the 3rd grade school year. Specifically, we wanted to address the extent to which the development of reading prosody adds to our ability to account for reading fluency and comprehension outcomes when children's word reading skills are included as well. To accomplish this, we added the TOWRE Sight Word Efficiency raw scores from times 1 and 2 to each of the outcome models tested previously. Essentially, if the prosodic features included in these models are in fact predictors of 3rd grade outcomes we expected that the significant relationships demonstrated in previous model tests would remain powerful despite the inclusion of the word reading skill variable. However, if either the pitch contour variable or the intrasentential pause variable (or both) no longer predicted the reading fluency and reading comprehension outcomes, or predicted these outcomes in a different way, we may reasonably conclude that the impact of automatic word reading skills accounts for the altered influence of these prosodic features on outcome measures of reading achievement and may therefore be the primary contributor to future reading performance as well. Our original intent was to include only paths reflecting both the direct and indirect effects of word reading skill on reading fluency and reading comprehension outcomes. That is, we planned to test these models without any paths between the prosodic variables and the TOWRE measurements. However, doing so resulted in generally poor model fit and the modification indices suggested that it was necessary to include paths between the pause and TOWRE variables. Consequently, we created additional paths (those which we felt had some justifiable reason for being included) in each model that examined the relationship between the number of pausal intrusions and the development of word reading speed and accuracy at the subsequent time point, and also between the number of pausal intrusions and word reading speed and accuracy at the same time point. The path diagrams for these models can be found in Figures 3.5 and 3.6.

Results of the Fluency as Predicted Outcome model (TOWRE included) indicated that the general structure of the skill development hypothesis between prosodic variables remained valid with all paths significant at the $p < .01$ level. As expected, performance on the TOWRE at time 1 significantly impacted performance at time 2 as well ($p < .01$). The cross-lagged relationship between the number of intrasentential pauses at time 1 and word reading speed and accuracy at time 2 was not significant; however, the relationship between the number of pausal intrusions at time 2 and word reading speed and accuracy at time 2 was significant at the $p < .05$ level. Furthermore, examination of the outcome relationships revealed several additional findings of definite theoretical importance. First, both the direct relationship between children's initial word reading ability and later oral reading fluency, and the indirect cumulative influence of the development of word reading speed and accuracy on oral reading fluency were found to be significant ($p < .05$). However, with the introduction of the TOWRE into the model, neither the indirect relationship between the cumulative effect of decreases in the number of pausal intrusions and later fluency nor the direct relationship between the number of pausal intrusions at time 1 and oral reading fluency at time 3 remained significant. Pitch contour, in contrast, emerged as a significant additional predictor of oral reading fluency in grade 3. Specifically, both the initial pitch contour measurement and the indirect cumulative effect of the development of appropriate pitch contour demonstrated significant relationships with the fluency achievement outcome measure ($p < .01$ for each). In previous analyses we suggested that decreases in pausing may initiate the development of pitch aspects of prosody, and further that pausing may be closely linked with word reading skills. Results of this model test suggest that this supposition may be correct and, in fact, word reading skill appears to account for the contribution of decreases in pausing with respect to fluency in grade 3 observed in the previous fluency outcome model, leaving pitch contour as a significant additional predictor. Path weights, standard errors, and *t* values for the Fluency as Predicted Outcome model (TOWRE included) are presented in Table 3.7. The overall fit of the model to the data was considered marginal to good, χ^2 (6) = 11.59 ($p = .072$), GFI = 0.97, NNFI = 0.97, CFI = 0.99, RMSEA = 0.098.

The findings for the Comprehension as Predicted Outcome model (TOWRE included) shared some similarities with those of the fluency model; however, important distinctions were evident. Again, the skill-development relationships between prosodic variables remained significant ($p < .01$) and performance on the TOWRE at time 1 significantly impacted performance at time 2 as well ($p < .01$). Though the initial TOWRE measurement was not a significant predictor of comprehension in grade 3, the cumulative effect of the development of word reading speed and accuracy over time significantly predicted comprehension skill at the end of the 3^{rd} grade school year ($p < .05$). These results are somewhat different from those of the fluency model in which both

TOWRE measurements demonstrated highly significant associations with later fluency achievement. Additional differences between outcome models with inclusion of the TOWRE concerned the relative contributions of the pause and pitch variables once word reading speed and accuracy were taken into account. Despite inclusion of the TOWRE, the indirect cumulative effect of decreases in the number of pausal intrusions over time demonstrated a significant influence on later comprehension skill ($p < .05$). The direct relationship between the initial number of pausal intrusions and comprehension skill, however, was not significant. Alternatively, whereas the direct relationship between the initial pitch contour measurement and comprehension skill in grade 3 was found to be significant at the $p < .01$ level, the indirect cumulative effect of the development of appropriate pitch contour did not predict later comprehension skill. Thus, inclusion of the TOWRE did not alter the previous findings from the Comprehension as Predicted Outcome model. Overall, the data were well fit by the model, χ^2 (6) = 9.75 (p = .14), GFI = 0.97, NNFI = 0.97, CFI = 0.99, RMSEA = 0.079. A summary of the path weights, standard errors, and *t* values for the Comprehension as Predicted Outcome (TOWRE Included) model can be found in Table 3.8.

In summary, it appears that the relative influence of the development of pause and pitch features varies according to the particular outcome, once word reading speed and accuracy is taken into account. Introduction of the word reading skill variable to the Fluency as Predicted Outcome model eliminated the significance of both the direct and indirect influence of pause on oral reading fluency. Instead, both the initial word reading skill level and the cumulative effect of the development of word reading skill were found to be significant predictors of later fluency achievement. Furthermore, with the addition of the TOWRE, pitch contour emerged as an additional significant predictor of oral reading fluency. Thus, in terms of fluency as an outcome, pitch contour appears to have additional influence whereas pause does not. Conversely, the results of the Comprehension as Predicted Outcome model were not affected by inclusion of the word reading skill variable. Specifically, both the initial pitch contour measurement and the cumulative effect of decreases in pausal intrusions remained as additional significant predictors of reading comprehension skill. The development of word reading skill simply served as yet another predictor as well.

Table 3.1
Descriptive Statistics for Full Developmental Study of Prosodic Reading

Prosody Feature	*N*	Minimum	Maximum	*M*	*SD*
Basic Declarative F_0 Change (Hz)					
Time 1	62	-78	12	-21	17
Time 2	69	-76	7	-22	19
Time 3	86	-82	-7	-30	16
Intrasentential Pause Total					
Time 1	62	0	21	6	6
Time 2	69	0	20	6	5
Time 3	86	0	12	3	3
Intersentential Pause Duration (ms)					
Time 1	62	309	2494	755	514
Time 2	69	113	2658	817	585
Time 3	86	78	1409	544	286
Phrase-final Comma Pause Duration (ms)					
Time 1	62	0	1630	391	385
Time 2	69	0	1721	410	375
Time 3	86	0	692	188	160
Child-Adult F_0 Contour Match					
Time 1	62	.041	.840	.530	.191
Time 2	69	-.024	.864	.549	.189
Time 3	86	.352	.864	.663	.101

Note. Time 1 = Spring of 1st grade, Time 2 = Fall of 2nd Grade, Time 3 = Spring of 2nd grade as outlined in the layout and assessment design for the full developmental study of prosodic reading. The differences in *N* at each time point are the result of missing data which most often was due to poor recordings. As the study progressed so did the technology and quality of the oral reading samples.

Table 3.2

Means, Standard Deviations, Skewness, and Kurtosis Values for Variables Used in Model Tests (n=92)

Variable	*M*	*SD*	*Skewness*	*Kurtosis*
Intrasentential Pause				
Time 1	7.457	5.894	0.824	-0.291
Time 2	2.837	3.100	1.129	0.397
Child-Adult F_0 Match				
Time 1	0.510	0.175	-0.462	-0.296
Time 2	0.660	0.099	-0.590	0.472
TOWRE-Sight Word Efficiency (raw score)				
Time 1	43.696	15.221	0.259	-0.148
Time 2	57.859	13.623	-0.288	-0.342
GORT-Fluency (sum score)				
Time 3	52.761	18.317	0.508	0.558
WIAT-Reading Comprehension (raw score)				
Time 3	21.967	4.819	-0.532	-0.532

Note. Time 1 = Spring of 1st grade, Time 2 = Spring of 2nd grade, and Time 3 = Spring of 3rd grade as outlined in the adjusted layout and assessment design for the full developmental study of prosodic text reading.

Table 3.3

Correlations between variables used in model tests

Variable	1	2	3	4	5	6	7	8
1. Intra Pause 1	--							
2. Intra Pause 2	.773**	--						
3. F_0 Match 1	-.384**	-.215	--					
4. F_0 Match 2	-.359**	-.313*	.363**	--				
5. TOWRE 1	-.774**	-.721**	.294*	.396**	--			
6. TOWRE 2	-.798**	-.745**	.379**	.383**	.864**	--		
7. GORT-Fluency	-.593**	-.627**	.270	.182	.770**	.770**	--	
8. WIAT-RC	-.365**	-.556**	.400**	.240*	.538**	.627**	.518**	--

Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions), TOWRE = Raw score on the Sight Word Efficiency Subtest of the Test of Word Reading Efficiency (TOWRE), GORT-Fluency = Sum Score on the Gray Oral Reading Tests, Fourth Edition (GORT-4), WIAT-RC = Raw score on the Reading Comprehension subtest of the Wechsler Individual Achievement Test (WIAT). The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1^{st} grade, 2 = Spring of 2^{nd} grade). *$p < .05$, **$p < .01$.

Table 3.4
Path Weights, Standard Errors, and t Values for Paths Models Examining the Relationship between Pause and Pitch Variables

Model/Path	Weight	SE	*t*
Independence:			
F_0 Match 1 → F_0 Match 2	0.39	0.098	3.98**
Intra Pause 1 → Intra Pause 2	0.81	0.062	13.06**
Skill Development:			
F_0 Match 1 → F_0 Match 2	0.28	0.101	2.76**
Intra Pause 1 → Intra Pause 2	0.81	0.062	13.06**
Intra Pause 1 → F_0 Match 2	-0.27	0.101	-2.66**
Reciprocal Effect:			
F_0 Match 1 → F_0 Match 2	0.28	0.101	2.76**
Intra Pause 1 → Intra Pause 2	0.84	0.067	12.49**
Intra Pause 1 → F_0 Match 2	-0.27	0.101	-2.66**
F_0 Match 1 → Intra Pause 2	0.07	0.067	1.04

Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions). The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). *$p < .05$, critical value = 1.96; **$p < .01$, critical value = 2.58.

Table 3.5
Path Weights, Standard Errors, and t Values for the Fluency as Predicted Outcome Model

Model/Path	Weight	SE	*t*
Skill Development Component:			
F_0 Match 1 → F_0 Match 2	0.28	0.101	2.76**
Intra Pause 1 → Intra Pause 2	0.81	0.062	13.06**
Intra Pause 1 → F_0 Match 2	-0.27	0.101	-2.66**
Fluency as Predicted Outcome:			
Direct Effects			
F_0 Match 1 → Fluency	0.22	0.081	2.73**
Intra Pause 1 → Fluency	-0.38	0.128	-2.98**
Indirect Effects			
F_0 Match 2 → Fluency	0.10	0.077	1.30
Intra Pause 2 → Fluency	-0.34	0.122	-2.78**

Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions), Fluency = The sum of the individual passage scores on the GORT-4, Form B administered during the spring of the 3rd grade school year. The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). *p < .05, critical value = 1.96; **p < .01, critical value = 2.58.

Table 3.6

Path Weights, Standard Errors, and t Values for the Comprehension as Predicted Outcome Model

Model/Path	Weight	SE	*t*
<u>Skill Development Component</u>:			
F_0 Match 1 → F_0 Match 2	0.28	0.101	2.76**
Intra Pause 1 → Intra Pause 2	0.81	0.062	13.06**
Intra Pause 1 → F_0 Match 2	-0.27	0.101	-2.66**
<u>Comprehension as Predicted Outcome</u>:			
Direct Effects			
F_0 Match 1 → WIAT-RC	0.35	0.086	4.05**
Intra Pause 1 → WIAT-RC	-0.01	0.250	-0.040
Indirect Effects			
F_0 Match 2 → WIAT-RC	0.02	0.095	0.21
Intra Pause 2 → WIAT-RC	-0.50	0.129	-3.88**

Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions), WIAT-RC = The raw number of points earned on the Reading Comprehension subtest of the Wechsler Individual Achievement Test administered during the spring of the 3rd grade school year. The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). **p* < .05, critical value = 1.96; ***p* < .01, critical value = 2.58.

Table 3.7

Path Weights, Standard Errors, and t Values for the Fluency as Predicted Outcome Model (TOWRE Included)

Model/Path	Weight	SE	*t*
Skill Development Component:			
F_0 Match 1 → F_0 Match 2	0.27	0.104	2.59**
Intra Pause 1 → Intra Pause 2	0.79	0.065	12.12**
Intra Pause 1 → F_0 Match 2	-0.29	0.104	-2.79**
TOWRE 1 → TOWRE 2	0.62	0.075	8.26**
Intra Pause 1 → TOWRE 2	-0.15	0.096	-1.56
Intra Pause 2 → TOWRE 2	-0.23	0.073	-3.13**
Fluency as Predicted Outcome:			
	Direct Effects		
F_0 Match 1 → Fluency	0.15	0.067	2.24*
Intra Pause 1 → Fluency	-0.13	0.083	-0.96
TOWRE 1 → Fluency	0.43	0.132	3.26**
	Indirect Effects		
F_0 Match 2 → Fluency	0.15	0.068	2.20*
Intra Pause 2 → Fluency	-0.16	0.104	-1.53
TOWRE 2 → Fluency	0.39	0.142	2.74**

Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions), TOWRE = Raw score on the Sight Word Efficiency Subtest of the Test of Word Reading Efficiency (TOWRE), Fluency = The sum of the individual passage scores on the GORT-4, Form B administered during the spring of the 3rd grade school year. The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). *p < .05, critical value = 1.96; **p < .01, critical value = 2.58.

Table 3.8

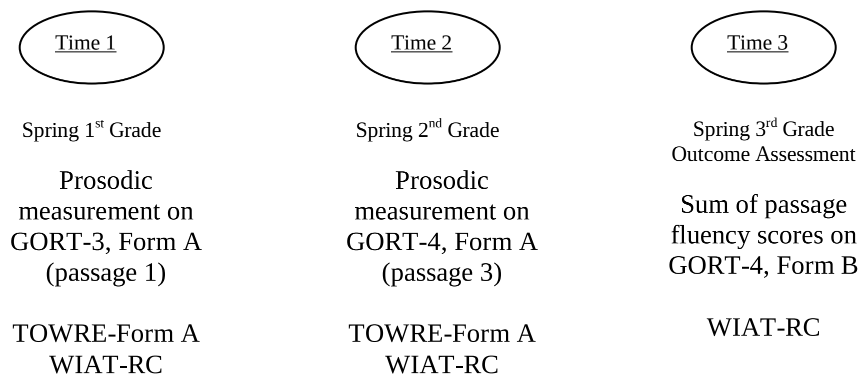
Path Weights, Standard Errors, and t Values for the Comprehension as Predicted Outcome Model (TOWRE Included)

Model/Path	Weight	SE	*t*
<u>Skill Development Component</u>:			
F_0 Match 1 → F_0 Match 2	0.28	0.104	2.68**
Intra Pause 1 → Intra Pause 2	0.79	0.065	12.11**
Intra Pause 1 → F_0 Match 2	-0.27	0.103	-2.62**
TOWRE 1 → TOWRE 2	0.64	0.073	8.82**
Intra Pause 1 → TOWRE 2	-0.12	0.092	-1.30
Intra Pause 2 → TOWRE 2	-0.24	0.068	-3.50**
<u>Comprehension as Predicted Outcome</u>:			
Direct Effects			
F_0 Match 1 → WIAT-RC	0.30	0.084	3.57**
Intra Pause 1 → WIAT-RC	-0.19	0.159	-1.19
TOWRE 1 → WIAT-RC	0.11	0.154	0.65
Indirect Effects			
F_0 Match 2 → WIAT-RC	0.05	0.093	0.54
Intra Pause 2 → WIAT-RC	-0.32	0.128	-2.50*
TOWRE 2 → WIAT-RC	0.37	0.184	2.01*

Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions), TOWRE = Raw score on the Sight Word Efficiency Subtest of the Test of Word Reading Efficiency (TOWRE). WIAT-RC = The raw number of points earned on the Reading Comprehension subtest of the Wechsler Individual Achievement Test administered during the spring of the 3rd grade school year. The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). *p < .05, critical value = 1.96; **p < .01, critical value = 2.58.

Figure 3.1

Adjusted layout and assessment design for the full developmental study of prosodic text reading (Phase 2)

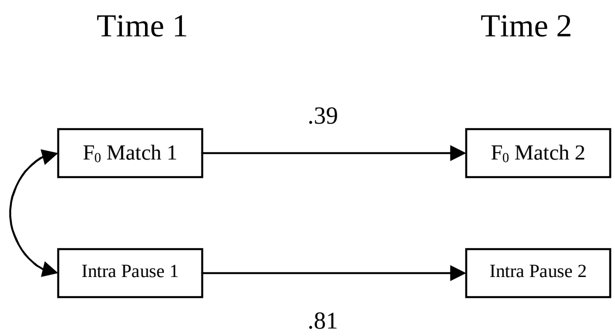


Note. Passage 1 on GORT-3, Form A and Passage 3 on GORT-4, Form A are identical.

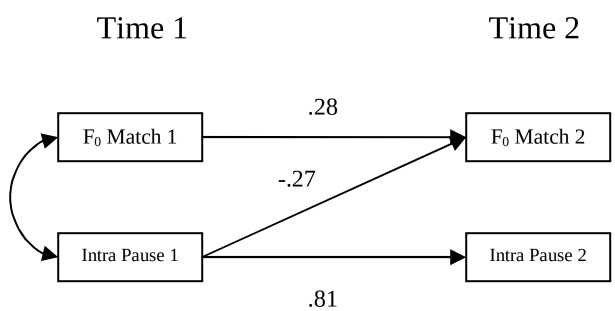
Figure 3.2

Path models tested to examine the relationship between intrasentential pause and child-adult F_0 contour match

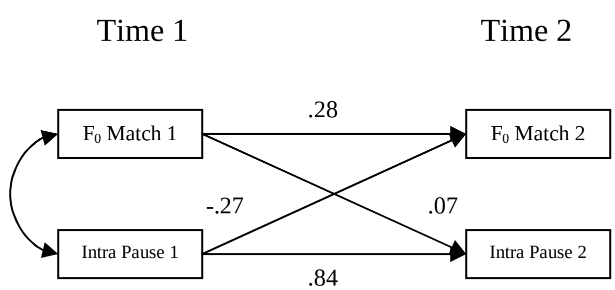
Model 1: Independence Model



Model 2: Skill Development Model



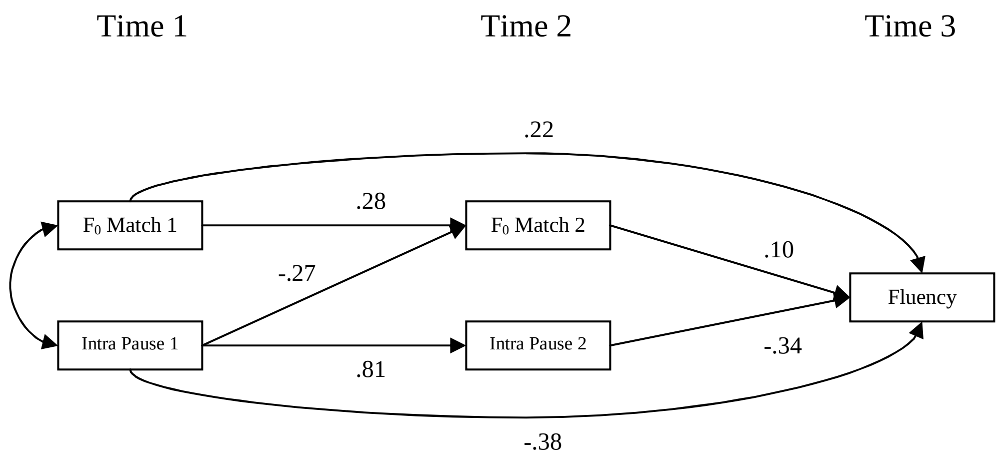
Model 3: Reciprocal Effect Model



Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions). The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade).

Figure 3.3
Path model tested to examine the relationship between intrasentential pause, child-adult F_0 match, and oral reading fluency

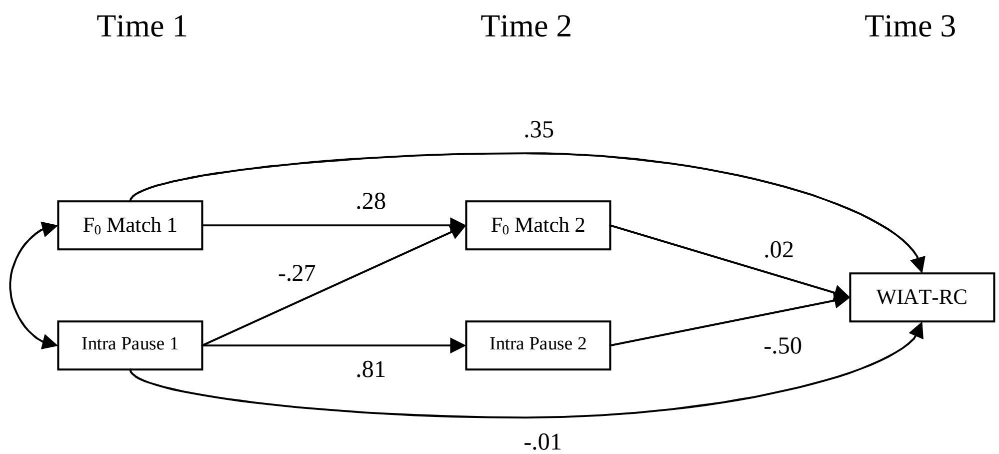
Model: Oral Reading Fluency as Predicted Outcome



Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions). The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). Fluency = The sum of the individual passage scores on the GORT-4, Form B administered during the spring of the 3rd grade school year.

Figure 3.4
Path model tested to examine the relationship between intrasentential pause, child-adult F_0 match, and reading comprehension

Model: Reading Comprehension as Predicted Outcome

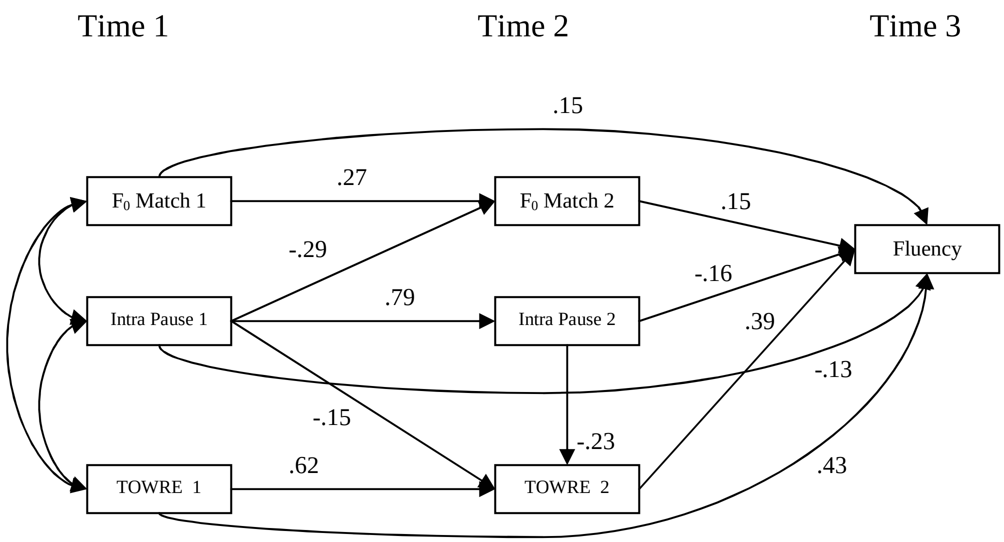


Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions). The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). WIAT-RC = The raw number of points earned on the Reading Comprehension subtest of the Wechsler Individual Achievement Test administered during the spring of the 3rd grade school year.

Figure 3.5

Path model tested to examine the relationship between intrasentential pause, child-adult F_0 match, and oral reading fluency with inclusion of the TOWRE

Model: Oral Reading Fluency as Predicted Outcome

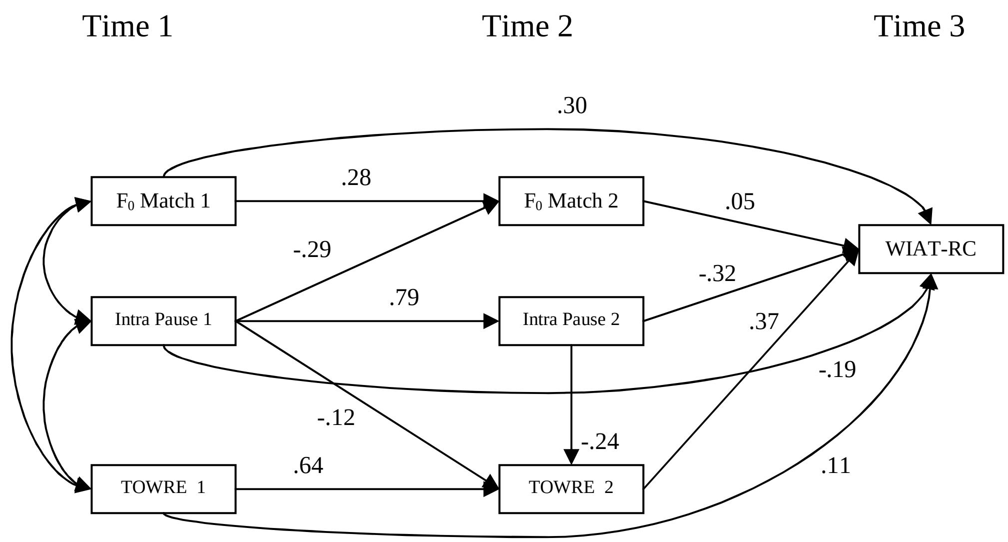


Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions), TOWRE = Raw score on the Sight Word Efficiency Subtest of the Test of Word Reading Efficiency, Form A (TOWRE). The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). Fluency = The sum of the individual passage scores on the GORT-4, Form B administered during the spring of the 3rd grade school year.

Figure 3.6

Path model tested to examine the relationship between intrasentential pause, child-adult F_0 match, and reading comprehension with inclusion of the TOWRE

Model: Reading Comprehension as Predicted Outcome



Note. F_0 Match = Child-adult F_0 contour match, Intra Pause = Total number of intrasentential pauses recorded during oral reading (pausal intrusions), TOWRE = Raw score on the Sight Word Efficiency Subtest of the Test of Word Reading Efficiency (TOWRE). WIAT-RC = The raw number of points earned on the Reading Comprehension subtest of the Wechsler Individual Achievement Test administered during the spring of the 3rd grade school year. The number after the variable name corresponds to the time point at which the measurement was taken (1 = Spring of 1st grade, 2 = Spring of 2nd grade). **p* < .05, critical value = 1.96; ***p* < .01, critical value = 2.58.

CHAPTER 4

GENERAL DISCUSSION

The present research represents the first known "long-term" developmental study of prosodic text reading in early elementary school children. In conducting this investigation our overall objectives were several and included the following: (1) to determine whether decreases in pausing over time serve a causal function for the development of larger pitch changes; (2) to determine the extent to which the growth of prosody during grades 1 and 2 is predictive of oral reading fluency in grade 3; (3) to determine whether the growth of reading prosody during grades 1 and 2 is predictive of comprehension skill in grade 3; and (4) to determine the extent to which the development of reading prosody adds to our ability to account for reading fluency and comprehension outcomes beyond word reading speed and accuracy.

Results of the preliminary phase of the study illustrated that various pause and pitch features of oral reading prosody appear to develop in a generally linear manner throughout the years of primary reading acquisition. Specifically, we found steady decreases in the number of pausal intrusions present in children's oral reading over time along with diminished durations in both intersentential and phrase-final comma pauses as well, all of which became increasingly consistent with "target" readings established in previous research (see Clay & Imlach, 1971; Dowhower, 1987; Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004). Furthermore, pitch features demonstrated steady developmental patterns as we observed a general increase in the magnitude of sentence-final pitch declination throughout our measurements and pitch contours that became increasingly similar to those of adults by the end of the 2nd grade. Each of these pitch renderings was consistent with those established in previous research as well (see Clay & Imlach, 1971; Dowhower, 1987; Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004). Such characterizations are of extreme theoretical importance given this type of systematic investigation had never been attempted previously and our understanding of prosodic development up to this point (particularly in terms of trend) has been purely speculative. Although Chall's (1996b) model of reading development suggests that prosodic text reading skills develop during the confirmation and fluency stage, which spans from the end of first grade to third grade, the absence of any empirical evidence in the research literature to support this assertion and elaborate on the explicit nature of prosodic development represents a critical theoretical gap in the achievement of skilled reading. However, findings from the present study provide support for Chall's model along with preliminary evidence regarding the particular growth patterns of selected prosodic features in children's oral reading.

In exploring the development of prosodic text reading in broader context we were interested in determining the specific relationship between the pause and pitch aspects of prosody. That is, our objective was to ascertain whether the pause and pitch variables selected for the model tests (intrasentential pause and child-adult F_0 contour match) develop simultaneously and independent of each other, whether decreases in the number of intrasentential pauses precedes the development of appropriate pitch contour (skill development), or whether a reciprocal relationship exists between the two variables (such that pause impacts the subsequent development of pitch and pitch impacts the subsequent development of pause). According to Chall (1996b) and Kuhn and Stahl (2003), prosodic text reading emerges in large part after children have solidified automatic decoding and word reading capabilities. Inherent in this theory is a skill development approach to understanding prosodic reading which suggests that as children's oral reading becomes increasingly fluent, they develop the ability to represent what is read in ways that imitate the tonal and rhythmic aspects of conversational speech. It follows that decreases in the number of pausal intrusions that result from the acquisition of automatic decoding and word reading skills might precede the development of adult-like pitch contour and perhaps serve a causal role in the emergence of pitch features as well. Model tests focused on three potential relationships concerning the development of pause and pitch features (independence, skill development, and reciprocal effect) with results ultimately supporting the skill development theory described above. Indeed, we found that while the initial pause and pitch measurements taken during the spring of 1st grade impact performance on their respective variables one year later (reflecting the "independence" aspect of the model tests), the initial pause measurement was found to influence subsequent development of pitch contour as well (evidence for the skill development relationship). Furthermore, our model tests failed to find a significant reciprocal relationship between the early development of pitch contour and later decreases in the number of pausal intrusions present during oral reading. Thus, support for the skill development theory was evident. Results suggested that the intrasentential pause measurement may, in fact, serve as a proxy for the achievement of automatic word reading skills and that the acquisition of these skills could serve an important role in the later development of pitch during oral reading.

Establishing the relationship between the development of pause and pitch prosodic features was a necessary first step prior to conducting the remaining model tests. Once the skill development relationship was identified we proceeded with additional tests that examined the extent to which the cumulative effects of the development of reading prosody predicted fluent reading and comprehension skills during the spring of the 3rd grade school year. Specifically, our purpose was to determine the respective influences of the development of pause and pitch on these reading outcome measures. Given the skill development relationship between the pause and pitch variables described above, we theorized that if prosody is more accurately characterized by the development of pitch contour (as opposed to decreases in pausing which instead might reflect the achievement of automatic word reading skills) then pitch might emerge as a more robust predictor of later fluency and comprehension skill. Such a belief was derived from the general consensus within the research literature that prosody is a higher-order component of fluency and the recognition that fluency and comprehension are strongly related. Results, however, were somewhat mixed with respect to this theory and the various reading achievement outcomes.

We found that while the indirect cumulative effect of decreases in the number of pausal intrusions during grades 1 and 2 predicted oral reading fluency in grade 3, the development of pitch contour was not a significant predictor of later fluency. Although this result runs contrary to theory, there are at least two possible explanations for this finding. First, it may be that decreases in pausing are simply a more powerful predictor of oral reading fluency when compared to the development of pitch contour. In fact, many traditional definitions of fluency consist only of rate and accuracy components, and the measure of fluency used in this study itself is basically an indicator of text reading speed and accuracy. Furthermore, our assumption that pitch might be a more robust predictor of oral reading fluency was actually based on previous research by Miller and Schwanenflugel (2006) and Schwanenflugel et al. (2004) in which pitch variables were found to be significant predictors of reading comprehension (not of fluency specifically). Given the strong correlation between fluency and comprehension skill, we hypothesized that pitch would serve as a predictor of fluency as an outcome in a similar way that it did for comprehension. Based on the results of this analysis, however, such an assumption may have been unfounded. Second, and somewhat related, is the notion that the development of pitch contour could emerge as a significant predictor of fluency in the future. Specifically, if the skill development hypothesis holds, it may be that the development of pitch features is somewhat delayed and that our study did not cover the appropriate time span to observe the desired effect. If prosodic measurements were continued during the 3rd grade school year it may have been possible to obtain different results. Still, we were able to find significant relationships between both the initial pause and pitch measurements and oral reading fluency. One final possibility to consider is that while both the initial number of pausal intrusions present in oral reading and the cumulative effect of decreases in pausing are significant predictors of fluency in grade 3, only the initial pitch contour measurement is an important predictor of later fluency. Thus, the developmental aspect of pitch contour may not be as meaningful in terms of predicting fluency achievement.

With respect to reading comprehension as an outcome the results of our model test yielded generally similar results to the fluency model explained above. Overall, we found that both the indirect cumulative effect of decreases in pausing throughout the course of development and the initial pitch contour measurement significantly predicted comprehension skill at the end of 3^{rd} grade. That the development of pitch contour did not impact reading comprehension, however, was somewhat surprising in light of previous research that demonstrated significant relationships between various pitch features (including child-adult F_0 contour match) and comprehension (see Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004). It is worth mentioning that neither the Miller and Schwanenflugel, nor the Schwanenflugel et al. studies examined prosody developmentally. Additional minor differences exist between the present research and the aforementioned studies as well (e.g., previous studies used somewhat older groups of children). Nevertheless, possible explanations for these findings are similar to those of the fluency outcome model. It may be that our study did not continue for the necessary duration to capture such a cumulative effect for pitch given a delay in the emergence of appropriate pitch contour due to the skill development relationship between pause and pitch variables. Alternatively, it is also possible that the initial pitch contour measurement, and not the developmental aspect of pitch, is simply a more influential predictor of later comprehension skill.

The final outcome model tests incorporated word reading speed and accuracy as an additional predictor to the original fluency and comprehension models. One possibility that remained to be explored was whether, and the extent to which, reading prosody added to our ability to account for reading achievement outcomes despite inclusion of the word reading skill variable. In the previous models we theorized that intrasentential pausing could be a proxy for automatic word reading skills. If true, it seemed likely that the effect of intrasentential pausing on fluency and comprehension would be usurped by the word reading speed and accuracy component in the models. However, if word reading skill impacts the separate influences of decreases in pausing and pitch changes similarly, then this might suggest that the same skills are responsible for each type of prosodic marker. Still, there was some uncertainty about the precise effect that word reading speed and accuracy would have on the pitch aspect of the model tests with respect to the outcome measures.

As anticipated, with the inclusion of word reading speed and accuracy to the fluency outcome model, results indicated that both the initial pause measurement and the indirect cumulative effect of decreases in pausing were no longer predictors of fluency in grade 3. The initial word reading skill measurement, along with the indirect cumulative effect of the development of word reading skills, however, were found to be strong predictors of later fluency achievement. Furthermore, pitch contour emerged as an additional predictor beyond word reading speed and accuracy. We found that both the initial pitch contour measurement and the indirect effect of the development of adult-like pitch contour over time impacted fluency performance at the end of the 3^{rd} grade school year. These findings were theoretically sound and generally consistent with the conclusions of previous research on reading prosody which suggested that pitch features may be the “true” indicators of prosodic text reading and, further, that the ability to produce appropriate pitch interpretations during oral reading predicts future performance in reading fluency and other reading-related skills as well (Clay & Imlach, 1971; Dowhower, 1987; Miller & Schwanenflugel, 2006; Schwanenflugel et al., 2004).

With respect to the reading comprehension outcome model, inclusion of word reading skill to the model test did not alter any of the previously established relationships between the pause and pitch variables and comprehension. We found that children's initial pitch contour measurement and the indirect cumulative effect of decreases in intrasentential pausing continued to impact later comprehension skill. Moreover, the indirect cumulative effect of the development of word reading automaticity predicted comprehension at the end of 3rd grade as well. That identical results were obtained for the two comprehension outcome models suggests that our original findings were accurate and that the inclusion of word reading skill does not appear to affect the prediction of comprehension in the same way as was evident in the fluency model. Thus, our obtained results seem to contradict certain conclusions put forth by previous research regarding the relationships between aspects of prosody and reading comprehension. Unlike the current study, Miller and Schwanenflugel (2006) and Schwanenflugel et al. (2004) failed to find an independent effect of pause structures on reading comprehension once reading speed and accuracy were taken into account. Our finding that decreases in the number of pausal intrusions over time was associated with comprehension skill is a unique result and it may be that other studies did not find such an effect because they did not examine pause structures developmentally. The combined findings from the aforementioned studies did support a general effect for pitch features on comprehension that was affirmed in the present study as well. However, our results suggest that the developmental aspect of pitch contour may not be as powerful as the initial child-adult F_0 match correlation in predicting later comprehension skill. Thus, it appears that children who are to be good comprehenders mark basic sentence structure through pitch from the onset of the achievement of reading fluency.

Summary and Implications for Educational Practice

We feel that the information presented here may have practical value for teachers who monitor their students' oral reading development. The results of this study are clear in demonstrating generally linear growth patterns for key pause and pitch prosodic features that are evident throughout grades 1 and 2. Furthermore, results also suggest that the early development of pause structures, specifically concerning observed decreases in the number pausal intrusions during oral reading, appears to impact subsequent development of appropriate pitch contour. An understanding of the processes involved in the development of prosodic text reading alone holds considerable value for educators particularly because such information increases awareness about what changes may occur in their students' reading expression over time. It follows that a more advanced understanding of the development of oral reading prosody may also foster more reasonable expectations regarding expressiveness aspects of children's reading and permit teachers to construct lessons that appropriately assist this progression as well. For example, teachers may wish to focus more resources on building students' automatic word reading skills in 1st grade and early 2nd grade. The likely result of which will be seen as a reduction in excessive pausing during oral reading that may ultimately spur the development of adult-like pitch profiles. In conjunction with modeling techniques, these combined efforts might be a rather effective means for producing prosodic readers. Most importantly, given the mounting evidence that the development of reading prosody has significant implications for fluency and comprehension outcomes, educators might reasonably conclude that children whose oral reading mimics the rhythmic and tonal qualities of language might be fully fluent and capable of understanding what they read.

Limitations of the Present Study and Future Research

The current investigation of the development of reading prosody has several issues that serve as limitations regarding the conclusions that can be drawn. The relatively small sample size available for the present research is considered the most significant and pervasive limitation of this developmental study of prosodic text reading in early elementary school children. Not only was a reduction in model complexity necessary given our limited sample size, but the strength of our conclusions based on the model tests may have suffered as well. Our original intent was to examine the development of various pause and pitch features at three separate measurement points throughout grades 1 and 2, along with an additional outcome measurement taken at the end of the 3rd grade school year. However, we were compelled to drop the fall of 2nd grade prosodic measurement (thus retaining the spring of 1st grade and spring of 2nd grade prosodic measurements) and select the single best indicators of pause and pitch (intrasentential pause total and child-adult F_0 contour match) for use in the model tests. Although the elimination of one of the prosodic measurement points and a reduction in the total number of pause and pitch variables examined in this study allowed us to conduct model tests that would produce reliable results, we were unable to evaluate models that included a more complete set of prosodic variables. Consequently, a more thorough understanding of the complex relationship between pause and pitch development and subsequent reading achievement outcomes could not be obtained from the analyses included in this study.

One additional limitation concerns the overall design of our study. It may be that a more appropriate longitudinal design would include prosodic measurements that extend beyond the 1^{st} and 2^{nd} grades, continuing into the 3^{rd} grade as well. Increasing the time duration of the longitudinal design might allow for a more accurate portrayal of the development of prosodic text reading and further make possible the detection of any delayed effects (particularly concerning pitch features) that might emerge throughout the years of primary reading acquisition.

Future research should directly address the limitations described above. Securing a large sample is essential for nearly all complex modeling techniques that may be appropriate for examining the development of prosodic text reading using multiple indicators of pause and pitch at each measurement point. When designing a “long-term” longitudinal study (whether it concerns the development of reading prosody or any other topic) one should make certain of the number of subjects that will be required for the proposed analyses while being mindful of the potential for significant sample attrition and any additional problems (e.g., technological difficulties associated with obtaining oral reading samples) that may reduce the total effective sample size as well. Prospective studies should also consider the theoretical importance of examining prosodic development into grade 3 and implement measurement designs that extend over a more inclusive period of reading development. One final consideration with respect to these recommendations for future research is the time required to conduct the individual prosodic measurements once the oral reading samples have been collected. With large samples and several waves of data collection the task of producing a usable data set could become daunting. As technology improves research involving the direct measurement of prosody may become less time intensive; however, at present significant time and personnel resources are required. This circumstance must be factored into future research endeavors.

References

Allington, R.L. (1983). Fluency: The neglected reading goal. *The Reading Teacher*, *36,* 556-561.

Anderson, R.C., Hiebert, E.F., Wilkinson, I.A.G., & Scott, J. (1985). *Becoming a nation of readers*. Champaign, IL: National Academy of Education and Center for the Studying of Reading.

Barron, R.W., & Pittenger, J.B. (1974). The effect of orthographic structure and lexical meaning on same-different-judgments. *Quarterly Journal of Experimental Psychology, 26*, 556-581.

Bates, E. (1976). The acquisition of pragmatic competence. *Journal of Child Language*, *1*, 227-281.

Bear, D.R. (1992). The prosody of oral reading and stages of word knowledge. In S. Templeton & D.R. Bear (Eds.), *Development of orthographic knowledge and the foundations of literacy: A memorial festschrift for Edmund H. Henderson* (pp. 137-189). Hillsdale, NJ: Erlbaum.

Berninger, V.W., Abbott, R.D., Billingsley, F., & Nagy, W. (2001). Processes underlying timing and fluency of reading: Efficiency, automaticity, coordination, and morphological awareness. In M.Wolf (Ed.), *Time*, *Fluency*, *and Dyslexia*. Timonium, MD: York Press.

Biemiller, A., & Levin, H. (1968). Studies of oral reading: pronounceability. In H. Levin, E.J. Gibson, & J.J. Gibson (Eds.), *The analysis of reading skill*. Ithaca, NY: Cornell University.

Boersma, Paul & Weenink, David (2004). Praat: doing phonetics by computer (Version

4.2.07) [Computer program]. Retrieved June 29, 2004, from http://www.praat.org/.

Carlson, K., Clifton, C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language, 34*, 383-398.

Carver, R.P. (1991). Using letter-naming speed to diagnose reading disability. *Remedial and Special Education, 12(5)*, 33-43.

Carver, R.P. (1997). Reading for one second, one minute, or one year from the perspective of rauding theory. *Scientific Studies of Reading, 1,* 3-43.

Cattell, J.M. (1886). The Time it Takes to See and Name Objects. *Mind*, vol. XI, 63-65.

Chafe, W. (1988). Punctuation and the prosody of written language. *Written Communication, 5*, 396-426.

Chall, J.S. (1996b). *Stages of Reading Development* (2nd ed.) Fort Worth, TX: Harcourt-Brace.

Chall, J. S., & Dale, E. (1995). Readability revisited: The new Dale-Call readability formula. Cambridge, MA: Brookline Books.

Chall, J.S., Jacobs, V., & Baldwin, L. (1990). *The Reading Crises*. Cambridge, MA: Harvard University Press.

Chomsky, C. (1978).When you still can't read in third grade after decoding, what? In S. J. Samuels (Ed.), *What research has to say about reading instruction* (pp.13-30). Newark, DE: International Reading Association.

Christophe, A., Dupoux, E., Bertoncini, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of the Acoustical Society of America, 95,* 1570-1580.

Christophe, A., Mehler, J., & Sebastian-Galles, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy, 2,* 385-394.

Clay, M.M., & Imlach, R.H. (1971). Juncture, pitch, and stress in reading. *Journal of Verbal Learning and Verbal Behavior*, *10*, 133-139.

Cooper, W.E., & Paccia-Cooper, J. (1980). Syntax and speech. Cambridge, MA: Harvard University Press.

Cowie, R., Douglas-Cowie, E., & Wichmann, A. (2002). Prosodic characteristics of skilled reading: Fluency and expressiveness in 8-10-year-old readers. *Language and Speech*, *45(1)*, 47-82.

Cromer, W. (1970). The difference model: A new explanation for some reading difficulties. *Journal of Educational Psychology*, *61*, 471-483.

Cruttenden, A. (1984). An experiment involving comprehension of intonation in children from 7 to 10. *Journal of Child Language*, *1*, 221-231.

Cruttenden, A. (1985). Intonation comprehension in 10-year-olds. *Journal of Child Language*, *12*, 643-661.

Cutler, A. & Swinney, D.A. (1987). Prosody and the development of Comprehension. *Journal of Child Language*, *14,* 145-167.

Daane, M.C., Campbell, J.R., Grigg, W.S., Goodman, M.J., and Oranje, A. (2005). *Fourth-Grade Students Reading Aloud: NAEP 2002 Special Study of Oral Reading* (NCES 2006-469). U.S. Department of Education. Institute of Education Sciences, National Center for Education Statistics. Washington, DC: Government Printing Office.

DeCarlo, L.T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, 2, 292-307.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39,* 1-38.

Doehring, D.G. (1976). Acquisition of rapid reading responses. *Monograph of the Society for Research in Child Development*, *165*(2).

Donahue, P.L., Voelkl, K.E., Campbell, J.R., & Mazzeo, J. (1999, March). *NAEP 1998 reading report card for the nation and states*. Washington, DC: Government Printing Office.

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

Dowhower, S.L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory into Practice, 30,* 158-164.

Echols, C.H., Crowhurst, M.J., & Childers, J.B. (1997). The perception of rhythmic units in speech by infants and adults. *Journal of Memory and Language, 36,* 202-225.

Ehri, L.C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading, 18,* 116-125.

Ehri, L.C. (1998).Grapheme-phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.). *Word recognition in beginning literacy* (pp. 3-40). Mahwah, NJ: Erlbaum Press.

Eichelman, W.H. (1970). Familiarity effects in the simultaneous matching task. *Journal of Experimental Psychology*, *86*, 275-282.

Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological*

Review, 100(2), 233-253.

Forster, K.I., & Chambers, S.M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12,* 627–635.

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.

Gerken, L., Jusczyk, P.W., & Mandel, D.R. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition, 51*, 237-265.

Gibson, E.J., Barron, R.I., & Garber, E.E. (1972). *The developmental convergence of meaning for words and pictures* (Appendix to Project No. 90046). Ithaca, NY: Cornell University.

Gibson, E.J., Osser, H., & Pick, A.D. (1963). A study of the development of grapheme-phoneme correspondences. *Journal of Verbal Learning and Verbal Behavior, 2,* 142-146.

Goldwave Digital Audio Editor (2004). St. Johns, New Foundland, Canada: Goldwave, Inc.

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language & Speech, 15,* 103-113.

Gough, P.B. (1996). How children learn to read and why they fail. *Annals of Dyslexia, 46*, 3-20.

Gout, A., Christophe, A., & Morgan, J.L. (2004). Phonological phrase boundaries constrain lexical access. II. Infant data. *Journal of Memory and Language, 51,* 548-567.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (eds.), *Syntax and semantics, Vol. 3: Speech acts* (pp. 225-242). New York: Seminar Press.

Hall, G.S. (1911). Educational Problems. Volumes 1 & 2. New York: D. Appleton & Company.

Harris, T.L., & Hodges, R.E. (Eds.) (1995). The Literacy Dictionary: The vocabulary of reading and writing. Newark, DE: International Reading Association.

Herman, P.A. (1985). The effect of repeated readings on reading rate, speech pauses, and word recognition accuracy. *Reading Research Quarterly*, *20*, 535-555.

Hirsh-Pasek, K., Kemler Nelson, D.G., Jusczyk, P.W., Wright-Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition, 26,* 269-286.

Houston, D.M., Santelmann, L.M., & Jusczyk, P.W. (2004). English-learning infants' segmentation of trisyllabic words from fluent speech. *Language and Cognitive Processes*, *19*, 97-136.

Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural Equation Modeling Concepts, Issues, and Applications* (pp. 158-176). Thousand Oaks: Sage Publications.

Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indices in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling*, *10*, 128-141.

Huey, E.B., (1968) *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press. (Original work published 1908).

Johns, J.L. (1993). *Informal reading inventories.* DeKalb, IL: Communitech International

Incorporated.

Johnson E.K., & Jusczyk, P.W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44,* 1-20.

Joreskog, K.G. & Sorbom, D. (2006). LISREL 8.8 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International.

Jusczyk, P.W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64,* 675-687.

Jusczyk, P.W., Hirsh-Pasek, K., Kemler-Nelson, D.G., Kennedy, L.J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24,* 252-293.

Jusczyk, P.W., Houston, D.M., & Newsome, M. (1996b). The beginning of word segmentation in English-learning infants. *Cognitive Psychology, 39,* 159-207.

Kame’enui, E.J., & Simmons, D.C., (2003). The DNA of reading fluency. *Scientific Studies of Reading*, *5*, 203–211.

Kame'enui, E.J., Simmons, D.C., Good, R.H., & Harn, B.A. (2001). The use of fluency-based measures in early identification and evaluation of intervention efficacy in schools. In M.Wolf (Ed.), *Time, Fluency, and Dyslexia*. New-York: York Press.

Karlin, A. (1985). Intonation in oral reading and reading comprehension. *Reading Horizons*, *25,* 169-175.

Kincaid, J.P., Fishburne, R.P, Rogers, R.L. & Chissom, B.S. (1975). Derivation of new readability formulas (automated readability index, Fog count and Flesch reading ease formula) for navy enlisted personnel. Research Branch Report 8-75. Naval Air Station, Memphis, TN.

Klatt, D. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics, 3,*129-140.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidences. *Journal of Acoustical Society of America, 59,* 1208-1221.

Kline, R.B. (2005). *Principles and practices of structural equation modeling* (2nd ed.). New York: Guilford Press.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. NT: Cambridge University Press.

Koriat, A., Greenberg, S.N., & Kreiner, H. (2002). The extraction of structure during reading: Evidence from reading prosody. *Memory & Cognition, 30(2),* 270-280.

Kraljic, T., & Brennan, S.E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or the addressee? *Cognitive Psychology, 50,* 194-231.

Kuhn, M.R. & Stahl, S.A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology, 95,* 3-21.

LaBerge, D., & Samuels, S.J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Levin, H., & Biemiller, A.J. (1968). Studies of oral reading: words versus pseudo words. In H. Levin, E.J. Gibson, & J.J. Gibson (Eds.), *The analysis of reading skill*. Ithaca, NY: Cornell University.

Logan, G.D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, *13*, 123-148.

Lyon, G.R. (1997). Learning to read: a call from research to action. Article adapted from statements made before the Committee on Education and the Workforce, U.S. House of Representatives, Washington, D.C., July 10, 1997.

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

Miller, J. & Schwanenflugel, P.J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology*, *98(4),* 839-853.

Morgan, J.L. (1996). Prosody and the roots of parsing. *Language & Cognitive Processes*, *11*, 69-106.

Morgan, J.L., & Saffran, J.R. (1995). Emerging integration of sequential and Suprasegmental information in preverbal speech segmentation. *Child Development, 66,* 911-936.

National Reading Panel. (2000). *Report of the Subgroups: National Reading Panel*. Washington, DC: National Institute of Child Health and Development.

Nazzi, T., Dilley, L.C., Shattuck-Hufnagel, S., & Jusczyk, P.W. (2005). English-learning infants' segmentation of verbs from fluent speech. *Language and Speech, 48(3),* 279-298.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordecht, Holland: Foris.

Nicholson, T. (1999). Reading comprehension processes. In G.B. Thompson, & T. Nicholson (Eds.), *Learning to read: Beyond phonics and whole language* (127-149). New York, NY, US: Teachers College Press.

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.

Perfetti, C.A., & Hogaboam, T. (1975). Relationship between single word decoding and reading comprehension skill. *Journal of Educational Psychology, 67,* 461-469.

Perfetti, C.A. (1985). *Reading Ability*. New York: Oxford University Press.

Prescott-Griffin, M.L., & Witherell, N.L. (2004). *Fluency in focus : comprehension strategies for all young readers*. Portsmouth, NH: Heinemann.

Pikulski, J.J. & Chard, D.J. (2005). Fluency: Bridge Between Decoding and Reading Comprehension. *The Reading Teacher*, 58(6), 510-519.

Pinnell, G.S., Pikulski, J.J., Wixson, K.K., Campbell, J.R., Gough, P.B., and Beatty, A.S. (1995). *L.istening to children read aloud.* U.S. Department of Education, National Center for Education Statistics, Washington, DC.

Psychological Corporation (1992). *Wechsler Individual Achievement Test*. San Antonio, TX: The Psychological Corporation.

Rasinski, T.V. (1990b). Investigating measures of reading fluency. *Educational Research Quarterly, 14(3),* 37-44.

Rasinski, T.V. (2004). Creating Fluent Readers. *Educational Leadership, 61*, 46-51.

Rasinski, T.V., Padak, N.D., McKeon, C.A., Wilfong, L.G., Friedauer, J.A., & Heim, P.

(2005). Is Reading Fluency a Key for Successful High School Reading? *Journal of Adolescent & Adult Literacy*, *49*(1), 22–27.

Reicher, G.M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*, 274-280.

Read, C., & Schreiber, P. (1982). Why short subjects are harder to find than long ones. In E. Wanner & L.R. Gleitman (Ed.), *Language acquisition: The state of the art* (pp. 78-101). New York: Cambridge University Press.

Richardson, V., Anders, P., Tidwell, D., & Lloyd, C. (1991). The relationship between teachers' beliefs and practices in reading comprehension instruction. *American Educational Research Journal, 28,* 559-586.

Samuels, S.J., Schermer, N., & Reinking, D. (1992). Reading fluency: Techniques for making decoding automatic. In S.J. Samuels & A.E. Farstrup (Eds.) *What research says about reading instruction* (2nd ed., pp. 124-144). Newark, DE: International Reading Association.

Schreiber, P.A. (1980). On the acquisition of reading fluency. *Journal of Reading Behavior, 12,* 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). San Diego, CA: Academic Press.

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.

Schumacker, R.E., & Lomax, R.G. (1996). *A beginner's guide to structural equation modeling.* Mahwah: Lawrence Erlbaum Associates.

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*, 119-129.

Selkirk, E.O. (1983). *Phonology and Syntax: The relation between sound and structure.* Cambridge, MA: MIT Press.

Shinn, M.R., Good, R.H., Knutson, N., Tilly, W.D., & Collins, V.L. (1992). Curriculum based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. *School Psychology Review, 21*, 459-479.

Snow, C.E., Burns, M.S., & Griffin, P. (1998). Preventing reading difficulties in young children. Washington, DC: National Academy Press.

Snow, D.P., & Coots, J.H. (1981). *Sentence perception in listening and reading* (Tech. Note 2-81/15). Los Alamitos, CA: Southwest Regional Laboratory.

Spache, G. (1953). A new readability formula for primary-grade reading materials. *The Elementary School Journal, 55*, 410-413.

Stanovich, K.E. (2000). Progress in Understanding Reading: Scientific Foundation and New Frontiers. New York: Guilford Press.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*, 142-164.

Tanaka, J.S. (1993). Multifaceted conceptions of fit in structural equation models. In K.A. Bollen & J. S. Long, (Ed.) *Testing structural equation models* (pp. 10-40). London: Sage Publications.

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

Turk, A.E., Jusczyk, P.W., & Gerken, L.A. (1995). Do English-learning infants use syllable weight to determine stress? *Language and Speech, 38,* 143-158.

Weiderholt, J.L., & Bryant, B.R. (1992). *Gray Oral Reading Tests, Third Edition*. Austin, TX: Pro-Ed.

Weiderholt, J.L., & Bryant, B.R. (2001). *Gray Oral Reading Tests, Fourth Edition*. Austin, TX: Pro-Ed.

Wheeler, D. D. (1970). Processes in word recognition. *Cognitive Psychology, 1*, 59-85.

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, *5(3),* 211–238.

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.

APPENDIX A

Plots Depicting the Development of Selected Prosodic Features During Grades 1 and 2

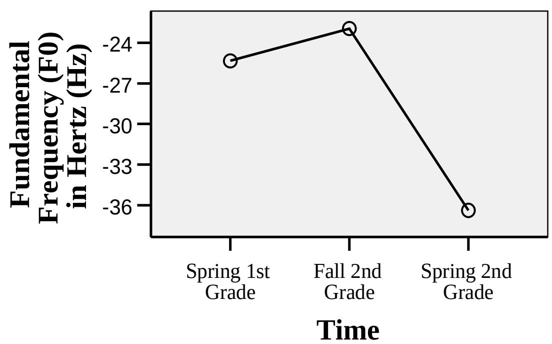


Figure A1. Basic Declarative Declination

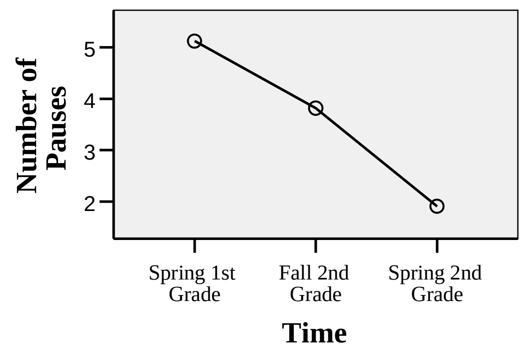


Figure A2. Intrasentential Pause

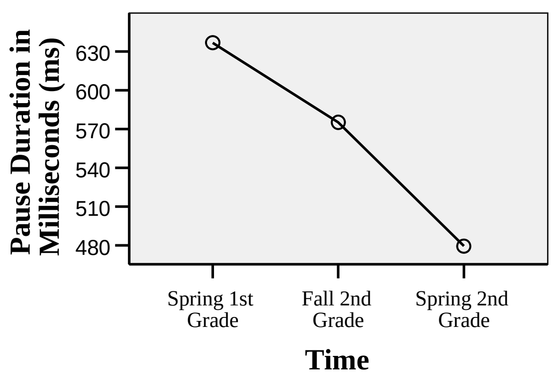


Figure A3. Intersentential Pause

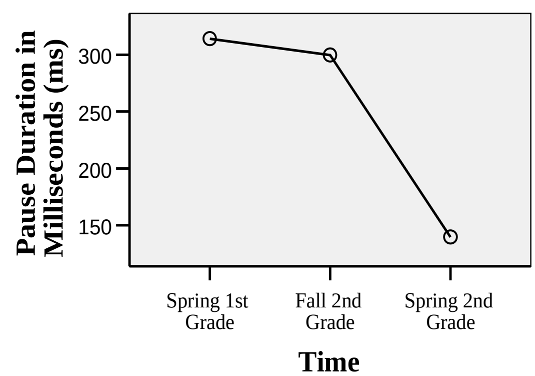


Figure A4. Comma Pause

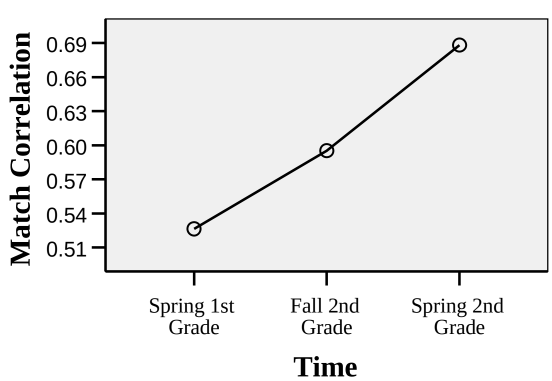


Figure A5. Child-Adult F_0 Match

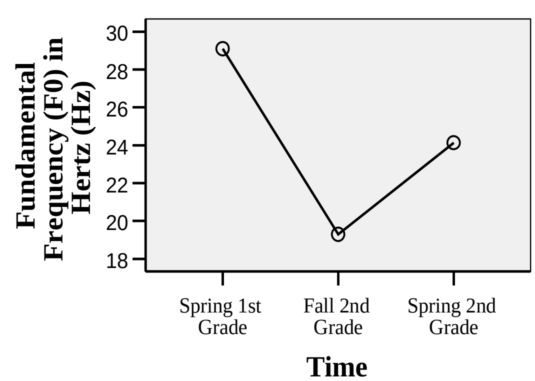


Figure A6. Yes-no Question

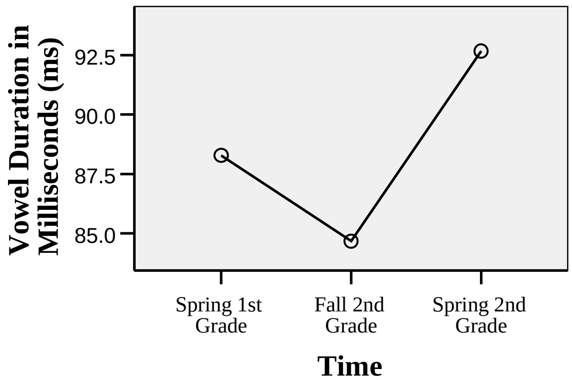


Figure A7. Phrase-final Vowel Elongation