I give permission for public access to my thesis and for any copying to be done at the discretion of the archives librarian and/or the College librarian.

Lianne Kaswer

May 2016

ABSTRACT

Reading comprehension may be the pinnacle skill in determining the success of a reader. It can be a difficult skill to master due to the amount of processes that affect its mastery. Prosodic fluency, the ability to appropriately produce prosody when reading aloud, and prosodic sensitivity, the ability to detect and utilize prosodic cues from audio sentences, have been shown to correspond with reading comprehension skill (Miller & Schwanenflugel, 2006; 2008). In the current paper we studied high school students with average and below average comprehension skill and measured their prosodic fluency from an imitation task, and their prosodic sensitivity utilizing a visual world paradigm. Sentences contained specific syntactic and semantic manipulations which elicit particular prosodic cues. We found that reading comprehension predicted one sentence type, ambiguous coordinate structures, in both experiments. This construct utilized prosodic phrase boundaries to signal height of attachment. Overall, better reading comprehension skill predicted better prosodic fluency and better accuracy performance. Reading comprehension may influence prosodic fluency and sensitivity, which in turn may affect differences in implicit prosody between poor and average readers.

The Relationship Between Reading Comprehension, Prosodic Fluency, and

Prosodic Sensitivity in High School Students

By

Lianne Kaswer

A Thesis

Presented to the Faculty of Mount Holyoke College

in partial fulfillment of the requirements

for the degree of Master of Arts

Department of Psychology and Education

Mount Holyoke College

South Hadley, Massachusetts

ACKNOWLEDGEMENTS

First and foremost I would like to thank my thesis advisor, Mara Breen, for her support and guidance throughout my time here at Mount Holyoke College. Thank you for giving me the opportunity to pursue this thesis project, as well as other opportunities which have challenged me and helped me grow as both a researcher and writer. You are an incredible advisor and I owe so much of my success to you. I would also like to thank my committee members, the chair of the department Kathy Binder, and Julie Van Dyke, for their patience and interest in my project.

Thank you to Cheryl Lee, Janelle Gagnon, Nicole Gilbert Cote, and Helin Unal, for all of your help and encouragement. All of you have taught me so much and helped me along this journey. Although no longer at Mount Holyoke College, thank you to the graduate students who came before me: Andrea Monson, Danielle Chapman, Danielle Denny, and Elizabeth O'Connor. Your friendship and support has meant so much to me and helped me through my work.

I thank my friends and family who have shown enthusiasm and support for all of my academic endeavors, regardless of whether or not they understood what it is that I have been working on!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Learning to read is no small feat. The various skills involved in that process are not underestimated by cognitive psychologists or educators. A successful reader must recognize graphemes, map phonemes correctly, grasp orthography, understand syntactic structures, decode words accurately and quickly, understand and produce fluent prosody, to semantically comprehend words, then sentences, then passages to complete books. As reading is a complex skill, the various aspects of reading and their relationships to each other are still being studied. More specifically, researchers are trying to understand which factors are the best predictors for reading comprehension ability, as it is one of the most important skills of the reading process to master.

Of those various factors that may contribute to the ultimate skill of reading is the presence of fluency. Colloquially, fluent readers simply sound good when reading aloud (Allington, 1983). Quantifying such a skill has proven difficult but necessary, as those readers who "sound good" are also more skilled at reading comprehension (Miller & Schwanenflugel, 2006; 2008). This fluency skill can incorporate several subskills, one being prosody, the focus of the current investigation. Prosody is a term which refers to a specific set of features present in spontaneous speech and those same features are also present when reading sentences aloud.

Prosody is concretely defined within its own field of study, however, it is not so easily defined when understanding its relationship to reading comprehension. Prosody is the "music" of spoken language, and contains the suprasegmental features, or the features of language that are not restricted by the individual segments of a word (i.e. phonemes, syllables etc.), they occur alongside those individual sounds and include rhythm, pitch, intensity, stress etc. (Lehiste, 1970). Pitch is commonly quantified as fundamental frequency (F0), which measures the fundamental frequency of waveform vibrations of speech in hertz (Hz) (Pierrehumbert, 1980). Rhythm can be understood as the regular temporal patterns of utterances which can be quantified by the timing of durational properties and emphasis (Nooteboom, 1997). Intensity can be thought of as loudness, which is usually measured in decibels (dB) (Vitz, 1972). The property of stress refers to either a word stress, in which a particular syllable(s) of a single word is stressed, or phrasal stress, which extends beyond one word to stressed words or phrases (Truckenbrodt, 2006). Stress can manifest as intensity, durational lengthening, or pitch accents. Any and all aspects of an utterance, from individual phonemes to complete phrases, carry prosodic information that can affect the overall timing, pitch and amplitude of the utterance and the sentence (Cutler, Oahan, & van Donselaar, 1997).

Prosody has been shown to reliably correspond to the syntax and semantics of a sentence (Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1997; Swerts,& Geluykens, 1994). Specific syntactic and semantic characteristics can correspond to specific prosodic cues. For example, declarative statements are characterized by the well-documented final F0 decline; the prosody of declarative statements conveys key syntactic information (Miller & Schwanenflugel, 2006).

Prosody can be used to “chunk” a sentence into meaningful phrases using intonation, pausing, etc., which aids with semantic integration (Kuhn, Schwanenflugel, Meisinger, Levy & Rasinski, 2010; Wagner & Watson, 2010). Prosodic cues such as the boundaries used to create phrases within a sentence can also be affected by phrase length; participants do take into account placing boundaries equally within a sentence (Fodor, 1998). Furthermore, these prosodic cues may not occur in isolation, several characteristics may be present to indicate the same syntactic information, like using pitch, pausing and duration to indicate a boundary. Prosody is a very complex component of spoken language, and therefore attempting to incorporate and measure it within the confines of a study may prove difficult.

Most of the research concerning prosody and reading comprehension refer to overt prosody, or prosody that is measured in spoken speech. Understanding prosody as it relates to reading comprehension presents a challenge, as there is no unifying operationalized definition. It has been described as reading with appropriate expression, and may or may not be grouped with other aspects of over-arching reading fluency, such as accuracy of reading (Dowhower, 1991; Kuhn et al., 2010). The features that constitute the criteria of “appropriate prosody” vary between researchers, which have created a disjointed understanding of prosody in reading comprehension. Moreover, the majority of research studying prosody and its effects on reading comprehension use elementary school children as participants, which has specific implications for the data. These implications may include the difficulty in generalizing the findings beyond young participants who are learning to read. The purpose of this current study was to operationalize the definition of prosodic fluency and investigate the relationship between prosody and reading comprehension for high school students.

Reading Fluency Skills and Reading Comprehension

Reading fluency, as it is more recently defined, is the ability to "read accurately at a quick pace with appropriate prosody" (Hudson, Pullen, Lane and Torgesen, 2009). Its effects on reading comprehension were and are still a broadly researched aspect of successful reading. Its relationship with reading comprehension is well documented (Benjamin & Schwanenflugel, 2010; Miller & Schwanenflugel, 2006; 2008; Binder et al., 2012). As research on reading fluency progressed, so did the development of including prosody as an aspect of reading fluency and as its own stand-alone component of successful reading.

Many processes may ultimately contribute to reading fluency, as Hudson, Pullen, Lane and Torgesen (2009) explored in their paper. They sought to add more depth to a definition of reading fluency which only measured the automaticity and accuracy in recognizing and pronouncing words. Within this model, they defined reading fluency itself as sight-word automaticity, decoding skills, orthographic knowledge and the integration of multiple cues. They found that reading comprehension, decoding fluency (how fast and accurately a participant can pronounce an unfamiliar word), and overall processing speed affects a reader's fluency. They argued that reading comprehension (containing skills such as vocabulary knowledge, passage context and metacognition) and reading fluency may have a reciprocal relationship, with one skill being advantageous to the other. However, their analysis of reading fluency may be too simplified as it is crucially missing the element of prosody, possibly because it may not be easily categorized to fit into a model. Prosody itself contains many characteristics that may have individual levels of influence on reading fluency levels, but that complexity does not warrant its exclusion from the relationship between reading fluency and reading comprehension.

Prosody is a component of language that has several measurable characteristics (e.g. pitch, duration, stress, pausing etc.) which can make it challenging to operationalize. It may be challenging in that researchers who study prosody must choose which prosodic features to measure (which characteristics are relevant to the study and what information they carry) and they must choose how to measure those features (such as by extracting quantitative acoustic features or using trained raters and an impressionistic scale) (Kuhn et al., 2010). Prior research has shown that appropriate prosody while reading is related with proficient reading comprehension (Miller & Schwanenflugel, 2006; 2008). Despite prosody's complexity, Kuhn et al., (2010) argued that because of this relationship between prosody and reading comprehension it is important to measure prosodic skill development alongside other aspects of reading fluency. Children who are developing reading skills are also learning how to use prosody effectively; which can be a challenge due to the complex nature of mapping prosody. Understanding the many types of information speakers and readers can glean from prosody is paramount to understand how it plays a role within reading comprehension. To effectively study that relationship, it is first necessary to know what previous psycholinguistic research has shown about the manifestation of prosodic characteristics and the information they can carry.

Psycholinguistic Research on Prosody

Psycholinguistic research has demonstrated that various aspects of prosody can indicate specific syntactic and semantic characteristics of a sentence; "prosody is syntax" (Wagner & Watson, 2010). Syntax, or the grammatical rules of a language, can be transmitted through prosody as well, such as through the realizations of boundaries. Moreover, English is a language in which the prosody of a sentence is influenced by its information structure; the content of the words, utterances and the overall sentence plays a role in determining its prosody (Féry & Kügler, 2008). The prosody of the information structure can be used to illustrate the semantics of a sentence, such as through prominences (Wagner & Watson, 2010) or by highlighting the focus or contrasting word of a sentence, which can change the overall meaning (Eady & Cooper, 1980; Ito & Speer, 2008; Pierrehumbert & Hirschberg, 1990). Prosodically grouping words within a sentence can convey important syntactic and semantic information as well, which facilitates overall sentence comprehension (Breen, Watson, & Gibson, 2011; Kjelgard & Speer, 1999; Snedeker & Trueswell, 2003). These various prosodic features have also been shown to be both produced by participants and utilized by participants when listening to audio stimuli. This dual modality indicates the significance of prosodic cues; they may be necessary for both accurate production and comprehension.

The scope of influence that prosody exerts on information structure was shown in the study of Eady and Cooper (1986) on focus words and sentence type. Participants were shown written stimuli which varied in terms of sentence type (statement or question) and focus (no focus, first or last sentence noun phrase) and were asked to produce them. Acoustic data, such as word duration and fundamental frequency (F0) contour were recorded. The questions served as priming stimuli in that they created focus expectations for the base declarative sentences as shown in (1)

(1) The ship is departing from France on Sunday

(1a) On what day is the ship departing from France?

(1b) What is departing from France on Sunday?

Whereas (1a) would elicit a focus on *Sunday* in the response sentence while (1b) would produce focus on *ship*. The researchers found that prosodic features were influenced by the presence of the focus word as well as the sentence type. The focus words were lengthened accordingly for emphasis, but the F0 contour for the focus words, as well as for the overall sentence, were influenced by the sentence type. Declarative statements would end in a F0 fall while questions would rise, with the same F0 pattern for each sentence contour was also found in the focus words for each respective sentence type. The overarching F0 contour observed in this study demonstrates the syntactic information which can be conveyed by prosody, and that it globally affects a sentence. The final pitch contour of a sentence, as seen through the marked difference between statements and questions, can be paramount in correctly conveying information and sentence type. Additionally, the results on the localization of duration indicate that singular words, if semantically significant to the sentence, will be prosodically emphasized. The results from the Eady and Cooper (1986) study show that prosody can convey information about both the sentence type and the words of importance within a sentence through duration and F0 contour.

Accenting is another method used to emphasize individual words of importance for comprehension. In an eye-tracking study, Ito and Speer (2008) instructed participants to view a grid with a complex arrangement of Christmas tree ornaments and were asked to decorate the tree according to audio instructions. These ornaments were described using a color adjective and the object noun of the ornament, with the color always preceding the object. Participants listened to instructions on what ornament to place on a tree, such as (2).

(2) Hang the green drum. Now hang the BLUE drum.

Stimuli in the contrastive trials included instructions which contrasted the color adjective, such as in example (2), or contrasted the object noun (e.g. *brown ball…brown ANGEL*). When relevantly contrastive information, such as *blue* from (2), was accented appropriately through pitch, participants were more likely to look earlier at relevant objects. When contrastive accents were placed on non-contrastive information for two instructions with no semantic connection, such as (3), participants were gardened-pathed and focused on the drum ornament instead of the angel ornaments as they anticipated only a color contrast, and not a contrast in object.

(3) Hang the green drum. Now hang the BLUE angel

These placements of pitch accents, along with the varied use of prosodic boundaries, can have a large impact on the overall syntactic and semantic understanding of sentences.

Prosodic boundaries are an important prosodic cue as they can convey different syntactic and semantic information. They are characterized by perceptual disjuncture between words or phrases that are realized through pausing, duration, pitch features, or a combination of features. They are placed at the end of a prosodic grouping, which can occur within a sentence or at the end of a sentence. Prosodic boundaries can be important for signaling both the syntactic relationship between words, and the semantic relationship between words, by prosodically grouping together words that are related (either syntactically or semantically) within a sentence for facilitated comprehension. Prosodically grouping words together to assist in comprehension can be seen through garden path sentences, such as

(4) The old man the boat

The initial parse of this sentence may incorrectly group *old* as an adjectival phrase with the noun phrase *man*. However, (4) can be prosodically disambiguated by instead grouping *the old* as the noun phrase while grouping *man* as a verb phrase with *the boat*. The disambiguating prosodic phrase boundary may take the form of increased duration of *old*, the word preceding the boundary between the noun and verb phrases. These phrase boundaries can be also realized by pauses between the phrase groups or through pitch in an intonational shift and are referred to as intonational phrase boundaries (IPh). The use of prosodic phrase boundaries in sentences are not uniformly placed according to one principle, rather, they may be modulated by several factors of the sentence, such as its syntax and semantic relationships between words.

The placement of prosodic boundaries can help disambiguate locally ambiguous clauses and ambiguous attachment. As noted in both Carlson (2009) and Frazier, Carlson, and Clifton (2006), subordinate clauses can cause temporary ambiguity if the syntactic boundary is not marked prosodically. Subordinate clauses are clauses which typically start with a conjunction and grammatically cannot stand alone; they are dependent on a main clause. This characteristic was exploited by Kjelgaard and Speer (1999), in that when a subordinate clause is recognized, it should prompt an expectation by the reader for it to be followed by a main clause. They created temporarily ambiguous subordinate phrases and studied how IPh (intonational phrase boundaries) affect comprehension. The participants listened to locally ambiguous fragments such as

(5) When Roger leaves the house…

(5a) is dark

(5b) it is dark

The initial fragment could lead to an interpretation in which *when Roger leaves* is a subordinate clause, or with *the house* as the direct object of *leaves*. Although the beginning fragment is identical in (5a) and (5b), there were both appropriate and inappropriate boundary tones and pauses associated with each; that classification differed depending on the target words, as certain syntactic structures elicit particular prosodic features. Subsequent analysis of their results showed that without an IPh, the phrase *the house* is taken as a direct object of *leaves*, which led the participants to an incorrect parse in (5a). However, with an IPh placed after *leaves* that interpretation is avoided and sentence (5a) is appropriately grouped and comprehended (Kjelgaard & Speer, 1999). Mismatching prosodic characteristics interfered with participants' ability to comprehend the sentences while appropriate prosodic features facilitated comprehension. These results provide support for the role that prosody plays in helping participants syntactically disambiguate and comprehend locally ambiguous sentences through prosodic boundaries.

Prosodic boundaries have also proven to be useful in disambiguating globally ambiguous sentences. In a study by Snedeker and Trueswell (2003), participants were assigned as either speakers or listeners; speakers had to successfully instruct the listeners to perform a task. The instructions were contextually ambiguous in certain conditions such as in (6):

(6) Tap the frog with the flower

Example sentence (6) can be construed as an instruction to tap a particular frog, in which *with the flower* is a modifier to indicate the correct frog to tap, or it instructs the listener to use *the flower* as the instrument of tapping. The correct interpretation of the stimulus cannot be construed without context; therefore the context was demonstrated to the speaker before they read the instructions to the listener. In those ambiguous conditions, the speaker's pausing and duration features disambiguated the instructions, and the listener paid attention to those cues. Specifically, when the direct object was lengthened in duration (*frog*) and a pause was inserted between the direct object and the preposition phrase (*with the flower*), this lead to the instrument interpretation. In the modifier interpretation, the verb was lengthened (*tap*) and a pause was inserted after; the direct object was prosodically grouped with the prepositional phrase. The prosodic phrase boundaries detected in the speaker's instructions differed between the two types of disambiguation, supporting the conclusion that prosodic boundaries are created in response to disambiguating syntactic attachment ambiguities and listeners utilize those cues to aid in their own disambiguation.

Prosodic boundaries are not only useful for syntactic disambiguation, but can aide in the comprehension of the semantics of the sentence as well. Breen, Watson, and Gibson (2011) studied the conditions of occurrence of prosodic boundaries within a sentence and whether the participants preferred to place boundaries at equal intervals (Fodor, 1998) or place them according to the semantic relationships between words (Wagner & Watson, 2010). They ran pairs of speakers and listeners with sentences in which the lengths of different utterances were manipulated. Speakers were asked to answer a comprehension question prior to producing the sentence for the listener. Prosodic boundaries were quantified as acoustic features of silence of phrase-final words and duration, and were measured through ToBI annotation. ToBI, the Tones and Breaks Indices system, is used to transcribe and annotate prosodic structures for acoustic data (Silverman, Beckman, Pitrelli, Ostendorf, Wightman, Price, Pierrehumbert, & Hirschberg, 1992). The researchers found that the participants more reliably produced the prosodic phrase boundary dependent on the semantic relationships between words. Due to the requirement of a comprehension question before production, the placement of prosodic phrase boundaries may have functioned as an indicator the speaker's semantic understanding of the sentence. In other words, after successful comprehension, the semantic relationships between words influenced the location of the phrase boundaries, strengthening the connection between prosodic cues to signal semantic structure.

Further studies have also shown evidence that the semantic relationships between words can be conveyed through prosodic boundaries. Dankovicová, Pigott, Wells, and Peppé (2004) studied how reliably children produce prosodic boundaries between compound nouns and a list of nouns, and how well adults can interpret their productions. Participants were shown pictures such as chocolate flavored biscuits (compound noun) or chocolate, biscuits, and honey (list). They measured the prosodic differences between conditions through the pause duration between the first and second nouns, and the final syllable duration of the first noun. Researchers found that overall participants produced prosodic boundaries appropriately according to the impressionistic ratings, but they noted a large amount of variation between the mean durations. The results support the idea that children can produce temporal boundaries with reliability, but fell short of producing completely adult-like prosody. This indicates that the children may recognize the semantic differences between the nouns in the conditions, but they cannot quite map the appropriate prosody. To (eventually) successfully map the correct prosodic cues onto semantic and syntactic structures, children will need to both recognize those structures and know what prosodic characteristics are needed.

The psycholinguistic research on the use of prosody in the comprehension of sentences is varied and sophisticated. The prosody at many levels of the sentence, from a syllable to a word, utterance, and phrase can carry important information needed for comprehension. The sophistication of that information can vary as well; ranging from a simple focus word that provides sentential context (Eady & Cooper, 1980), to prosodic boundaries which can disambiguate globally ambiguous stimuli (Snedeker & Trueswell, 2003). The prosodic cues used to signal that information are not always standard, with some cues, such as F0 change, manifesting in a range of possible prosodic characteristics (Eady & Cooper, 1980; Ito & Speer, 2008; Kjelgaard & Speer, 1999). By better understanding how the location and the prosodic characteristics used relate to specific aspects of sentence structure, we can more stringently measure the prosodic ability of readers. This allows us to better pinpoint where in the sentence a reader should use prosodic cues, and how those cues facilitate the comprehension of the sentence.

Reading Comprehension Skills and Prosodic Fluency

Understanding the relationship between prosody and reading comprehension involves not only researchers in the fields of linguistic and psychology, but also those in the field of education. Prosody, although not always outright identified, is learned alongside other reading skills; therefore researching the learning process of prosody can illuminate its role in reading (Dowhower, 1991). Educators who worked first hand with students developing their reading skills have recognized the importance of prosodic skills for successful readers.

One such researcher, Schreiber (1991), compiled the findings of two of his prior studies to overview the importance of prosody for children learning to read. The author posited that reading successfully requires readers to appropriately "chunk" phrases to structure the sentence for comprehension. Schreiber argues that when children are learning to read they are also learning to structure, and rely heavily on prosody to cue them. In his 1982 study, participants were asked to identify a phrasal target from a sentence which contained appropriate or inappropriate prosody. The child participants had lower rates of success than adults in identifying the phrasal targets when the prosody of the stimuli was inappropriately marked. These results may indicate that children rely more heavily on prosody to cue a meaningful phrase. Schreiber's 1987 study utilized the next-word paradigm in which he presented sentences, either with or without appropriate prosody, to both children and adults, and asked them to identify the word which immediately followed a chosen probe word in the sentence. Children's response times slowed more than the adults' when the probe word and its following word were presented in a prosodically inappropriate sentence, as this may have impeded the children's ability to successfully group the phrases within a sentence. Effective prosody was recognized as present in already successful and fluent readers, but Schreiber's work supports the notion that prosodic cues are markedly important to children while they are learning to full comprehend sentences. However, Schreiber recognized that children must be taught how to recognize the prosody of a written sentence, as it has a relationship with both reading fluency as well as the comprehension of a sentence.

Dowhower (1991), like Schreiber, recognized the importance of teaching prosodic skills for children learning to read. Through the review of prior research, Dowhower created an outline of specific prosodic markers that indicate reading with expression. These included pausal intrusions, the length of phrases, appropriate phrasing, phrase-final lengthening, stress, final intonational contours. Dowhower argued that although the relationship between prosody and reading comprehension was undefined, there is a possible connection between prosody and its effects on comprehension, noting that inappropriate prosody could adversely affect comprehension. It is clear that both Schreiber and Dowhower recognized and supported the idea that prosodic fluency may be crucially linked to reading, and further research studying that relationship followed in due time.

Later empirical research on the relationship between prosody, fluency, and reading comprehension is varied, in terms of methodology as well as terminology. Despite their differences, the research supports a relationship between prosody and reading comprehension, especially in the use of prosodic skill as a predictor for reading comprehension. Though at times diverse, these studies have furthered the understanding of this relationship and motivated our current study and definition of prosodic fluency.

In the earlier of several studies concerning prosody and reading comprehension, Schwanenflugel, Hamilton, Kuhn, Wisenbaker, and Stahl (2004) measured prosody, decoding speed and their relationship to reading comprehension ability in 2nd and 3rd graders. Decoding speed measured the number of words participants pronounced accurately in a 45 second period. Prosody was measured through the characteristics of inter and intrasentential pausing and F0 contours while reading only declarative sentences. The researchers found the strongest relationship between faster decoding skills, better reading comprehension ability, and more adult like F0 contour. There were, however, limitations. The pausing measures utilized for prosody may be mechanically more related to decoding skills than prosodic skills, in that readers who are faster and more accurate at recognizing words will make fewer errors that could force them to slow down, pause, or stumble while reading, which would be measured as inappropriate prosody. In other words, the prosodic measures used may not accurately reflect prosody. Additionally, the homogeneity of the reading materials is problematic, as psycholinguistic research has shown that prosodic markers are utilized in a diverse range of stimuli. Limiting the reading material to one sentence type constrains the experiment's ability to elicit a broad range of the prosodic features available, and so the results only capture a portion of what prosody is used for. Specifying more distinct prosodic characteristics, as well as differentiating the reading stimuli would paint a much richer and more inclusive picture of prosody's role in reading comprehension.

In a later study, Miller and Schwanenflugel (2006) improved upon the earlier Schwanenflugel et al., (2004) study by measuring specific prosodic features produced by school-age children reading several various syntactically complex sentences. This passage contained declaratives, basic quotatives, wh-questions, yes-no questions, complex adjectival phrase commas, and phrase-final commas, a very diverse stimulus set as compared to the 2004 study. Each of these sentence types included a predicted prosodic contour (based on the structures within the sentence); these contours specified which prosodic feature(s) would manifest and their location. Miller and Schwanenflugel found that reading skill level was correlated with adult-like pitch changes and pauses, indicating that children who changed pitch and paused in the appropriate context like prosodically fluent adults were more likely to be more proficient readers. By incorporating sentence types and specifying the prosodic structures associated with them, this study more accurately measured the presence of appropriate prosody. As the prior psycholinguistic research has shown, the prosody present in sentence varies in both the types of information it signals and the characteristics used in signaling them. The Miller and Schwanenflugel (2006) study quantified appropriate prosody and therefore was better able to identify instances of prosodic fluency and how it related to reading comprehension ability.

Although Miller and Schwanenflugel (2006) captured more of the complexities of prosody and reading comprehension than Schwanenflugel et al.'s 2004 study, they cannot speak to the temporal development of prosody. Longitudinal studies, however, are better equipped to explore how prosody develops in conjunction with measuring reading ability. Miller and Schwanenflugel (2008) investigated the relationship between prosodic fluency and reading comprehension by testing how specific prosodic markers impact reading comprehension and oral reading prosody. The researchers tested the oral fluency, word reading skills, reading comprehension and prosodic skills of children over the course of a year. Participants read a passage with various types of sentences. Prosodic skill was evaluated by measuring intersentential pause durations, phrase-final comma duration, pausal intrusion duration, sentence final pitch declination, and intonational contour. Results showed that prosodic markers, especially pausal intrusion and pitch contour, were the strongest predictors of gains in fluency as well as reading comprehension. They posited that as students gain better fluent reading skills (i.e. more accurate and faster word recognition) they are able to better produce adult like prosody. This model supports the idea that not only does prosodic skill predict reading comprehension ability, but that certain prosodic features are stronger indicators and predictors of reading comprehension than others.

The prior studies measured the relationship between prosody and reading comprehension as dependent on the sentence type and age of the participants. One aspect of reading that was not manipulated within those studies was the difficulty of the texts. As students continue to develop and hone their reading skills, over time, the texts given to them will progress in difficulty. Benjamin and Schwanenflugel (2010) studied the effects of text difficulty and how it may interact with prosodic reading fluency. Students were given passages of texts which advanced in difficulty and were asked to read them aloud. Researchers measured the prosodic features of sentence-final F0 change, intonation contour, intrasentential pausing, and ungrammatical pausing. They found that for the more difficult texts, prosody was more strongly related to the other aspects of reading fluency (reading rate and accuracy). The analyses also revealed that prosodic ability predicted reading comprehension scores independently of the other reading measures for the difficult texts, but not for the easy passages. This particular finding is interesting, in that it provides possible evidence for how prosody is used for comprehension after basic reading skills are mastered, which may extend to older students and adults.

Prosody is one of the various ways a reader can comprehend a sentence (other methods include vocabulary, phonological and orthographic awareness, verbal and nonverbal IQ, etc.), and according to the previous studies, prosodic skill is strongly related to reading comprehension ability. Other studies have shown that children learning to read were more likely to use prosody for comprehension purposes than successful adult readers (Schreiber, 1991). Children's reliance on prosody may indicate that prosody is a well ingrained tool used to signal varying structures for comprehension (Dankovicová et al., 2004). However, if a successful reader is presented with a challenging textual passage, they may then revert to using prosodic features to aid in comprehension when other methods are not enough. This may be particularly important for older readers who may be faced with more difficult text on a regular basis; for those challenging texts using prosody is an effective strategy. However, if an adult reader progressed beyond primary school with difficulties in their prosodic skills, these harder texts may pose an even greater challenge to them.

Reading Comprehension Skills and Prosodic Fluency with Adolescent or Adult Participants

Some of the work on prosody and reading comprehension has studied adolescent and adult participants. Participants past elementary school age may be less likely to be disfluent unlike younger participants who are learning to read (i.e. inappropriate pauses) as this disfluency can affect the prosodic characteristics being measured. As children are learning to read, they are learning several skills in parallel, and therefore they may be by happenstance more skilled in some areas than others. This may lead to labelling temporary disfluency as a possible skill deficiency when it may be more indicative of the priorities in their education curriculum at the time (Lai, Benjamin, Schwanenflugel & Kuhn, 2014). The amount of skills they are learning in tandem may also produce more inappropriate pauses, which do not necessarily indicate the presence of a prosodic break but perhaps are attempts to allow the reader a break to process what was read. Additionally, older participants are more likely to realize a broader range of prosodic characteristics (Wells, Peppe, & Goulandris, 2004).

There is some evidence that fluency has been shown to predict reading comprehension ability in older readers (Binder et al., 2013; Paige et al., 2014; Tighe & Schatschneider, 2014). For example, the meta-analysis conducted by Tighe and Schatschneider (2014) studied various predictors of reading comprehension skills, and explored the influence of each on reading comprehension scores of struggling adult readers. Of the 16 studies selected for inclusion, the results indicated that of the 10 components related to reading comprehension, fluency had a strong relation with reading comprehension scores. The strength of fluency in predicting reading comprehension skills is important in both children and low-literacy adults (defined as adults with decoding skill deficits), and its role was consistently found in the studies overall. However, Tighe and Schatschneider did not separate any components of fluency; as long as a study was measuring fluency and labeled it as such, the study could be included. This allowed for a diverse set of definitions of fluency to be included, which consequently confuses the results of their analysis. Even though this study did not specifically analyze prosody's role in reading comprehension over several studies, it still demonstrated the overall role of fluency in predicting reading comprehension in skill in adult participants.

Other studies have shown similar relationships specifically between prosodic fluency and reading comprehension in adolescent students. Paige, Rasinski, Magpuri-Lavell, and Smith (2014) explored this relationship in ninth grade participants by measuring their prosodic fluency with a scale which measured prosodic fluency according to broad qualitative characteristics. This scale (Multi-Dimensional Fluency Scale) rated the overall prosody of each participant by the appropriateness of expression and volume, smoothness, phrasing, and conversational pacing (Zutell & Rasinski, 1991). Their results indicated that prosody accounted for a significant amount of variance between the reading comprehension scores; however, as an impressionistic scale was used to measure prosody, the conclusions the researchers can draw were limited. The results from Paige et al., refer to very broad elements of prosodic fluency, and because of that breadth any differences between the prosody due to different sentence types, syntactic or semantic structures, and prosodic characteristics themselves are lost. Utilizing quantitative measures instead of an impressionistic scale would give more support to their findings.

Binder, Tighe, Jiang, and Kaftanski, (2013) sought to understand how struggling adult readers utilize prosodic fluency by studying different acoustic measures of prosody along with other traditional predictors of reading comprehension. The researchers utilized quantitative measures to explore prosodic fluency. Specifically, they explored these relationships with both skilled and low-skilled literacy adults enrolled in an educational course. Researchers measured phonological awareness, decoding, and word recognition, reading rate and prosody; prosody was measured through five types of pausal indicators and three types of pitch indicators. The pausing behavior included pauses after four chosen words, three complex adjectival commas, three phrase final commas, four sentence final pauses, and four simple quotatives. The pitch indicators consisted of final pitch rises for wh- and yes/no questions, and final pitch falls for declarative statements. Overall their results for low-literacy adults paralleled those of children, such as those from Miller and Schwanenflugel (2008). The punctuation marks within a sentence elicited a pause from both skilled and low skilled readers, and although there was a pitch declination present at the end of declarative sentences for both skilled groups, generally low-literacy skilled adults had smaller pitch variation than skilled counterparts. Pausing rates and lengths were related to decoding and word recognition abilities, and they correlated with reading comprehension as well. Binder et al. explained the increase of pausing and intrusions and its relation to reading rate and word recognition may be that when a low-literacy skilled reader cannot easily comprehend a sentence, the reading rate is slowed down to allow for more time for comprehension. This rate decrease, although helpful for comprehension, may increase the presence of pausal intrusions and inappropriate pauses, which could inadvertently be attributed to poor prosodic skills. Our study expands upon some of the methodologies used in Binder et al., in addition to studying a different population with different reading skill levels.

Defining Prosodic Fluency

The psycholinguistic research on prosody reveals it is a multifaceted and integral aspect of speech. The research measuring prosody as an aspect of successful reading also conveys that message, albeit through the variety of its definitions and measurements. The previous definitions of prosody and prosodic skill fail to capture what it means to prosodically fluent in terms of expressing prosody in relation to reading. We define prosodic fluency as the ability to appropriately produce prosodic characteristics as determined by the syntactic and semantic context. In our study those particular sentence structures and contexts have clearly defined prosodic characteristics, allowing us to understand what prosodic characteristics are suitable in which stimuli and when in the sentence. Furthermore, the quantification of prosodic fluency makes our study's indicators of prosodic fluency universal; the same measures can be applied with the same type of sentences to test for prosodic fluency. Within the previous literature, quantitative approaches were used as well as impressionistic scales, creating differing ideas of what constituted as “good prosody”. Measuring prosody can be better standardized through the use of quantitative measures and specifying the location of specific prosodic characteristics; this methodology can be more easily replicated. This type of quantification is not limited to only prosodic production; it can also be useful for measuring (and standardizing) how participants understand prosodic cues.

Reading Comprehension Skills and Prosodic Sensitivity

As evidenced with prior psycholinguistic research, prosody is utilized by both speakers and listeners: participants both produce prosody and detect prosodic cues to aid in comprehension. Research on prosody and reading comprehension has taken the same approach by studying prosodic sensitivity, or the ability to accurately detect and understand prosodic cues. Although many studies of prosodic fluency and reading comprehension utilize production data from participants, that alone cannot illustrate a full understanding of that relationship. Researchers use listening comprehension tasks as a measurement of understanding prosody, in this case prosodic sensitivity, and its relationship with reading comprehension. The prosodic characteristics present in the audio stimuli can be controlled by manipulating the structures (both semantic and syntactic) within the sentences used for the stimuli. The results from such stimuli will more directly reflect how prosodic sensitivity, or the sensitivity to specific characteristics of it, can affect comprehension.

Prosodic sensitivity studies can fall prey to the same limitations of prosodic fluency studies: both types of studies should strive for diverse sentential stimuli to elicit a wide range of prosodic characteristics. Although limited in its prosodic sensitivity scope, studies such as Clin, Wade-Woolley, and Heggie (2009) can still inform us about prosodic sensitivity through its particular characteristic of choice. Clin et al., (2009) studied the relationship between prosodic sensitivity, morphological awareness and reading ability using listening comprehension tasks. The researchers measured morphological awareness (Carlisle, 1988) by asking students to listen to a word and produce a new word based off of its stem (e.g. *election* from *elect*). The responses were categorized depending on the relationship between the response and the stem word (no change, a phonemic change, a stress change, or a stress and phonemic change). Prosodic sensitivity was measured using two stress tasks: the DeeDEE task and the Stress Contour Discrimination Task.

The DeeDEE task assessed the sensitivity to the prosodic structure at the phrasal level, by first having participants listen to a familiar phrase (such as a book title). The participants then heard two phrases with no phonemic content but an intact stress pattern and had to match the nonsense phrase to the original phrase. The other stress task, the Stress Contour Discrimination Task, measured the overarching stress pattern of a sentence. Participants had to determine if a subsequent sentence with no phonemic content matched the original sentence. Reading ability was measured through various skills, such as speed, accuracy, decoding etc. Clin et al., (2009) found that after accounting for phonological awareness, which is a strong predictor of reading ability (Hintze, Ryan, Stoner, 2003), prosodic sensitivity was still a significant predictor of reading ability.

This study strengthens the evidence for prosodic sensitivity's role in reading comprehension. However, because stress was the only prosodic feature tested in this experiment, the conclusions about prosodic sensitivity and reading comprehension are limited. Other studies (Miller & Schwanenflugel, 2006; 2008), have noted that different characteristics of prosody, such as duration or F0, differed in their ability to predict reading comprehension skill. Without testing for a wider range of prosodic characteristics, this study cannot detect differences between those characteristics and how they may interact with reading comprehension ability. It is important to not only include a variety of sentence types, but to ensure the prosodic characteristics of interest are varied as well. Those same differences that were found in production studies may very well hold true in prosodic sensitivity research.

Indeed, other researchers have attempted to isolate several specific aspects of prosody for exploration in prosodic sensitivity. Holliman, Williams, Mundy, Wood, Hart, and Waldron (2014) explored prosodic sensitivity and reading ability in a multi-component model, which included stress, intonation and timing on the word to word level, phrase level, as well as overall in a given sentence. Children listened to phrases or full sentences, which were either clearly pronounced with phonemic content, or unclear with no phonemic content but with intact prosodic contour. For each trial, participants were given pictures, and they had to match the audio utterance with the correct visual representation. The performance on the matching task with the stimuli lacking phonemic content was strongly correlated with reading comprehension ability, and their analysis pointed to possible differences in accuracy between intonation, stress and timing. Based on their exploratory factor analysis, the type of prosodic characteristic (intonation, stress, and timing) appeared to be more influential on the variance the factor explained than the word level; again, indicating that further testing could more concretely identify how each of those features of prosodic senstivity may contribute differently to comprehension.

Other listening comprehension studies, such as Whalley and Hansen (2006), have taken further steps in finding evidence for differences in prosodic characteristics and reading ability by empirically measuring specific characteristics of prosodic sensitivity. Two tasks were used to measure prosodic sensitivity: a DEEdee task, and the compound noun task which measured intonation, pauses, and stress at the word level by asking participants to discriminate between compound nouns (e.g. highchair) and noun phrases (e.g. high chair). Results indicated that performance on the DEEdee task predicted reading comprehension ability while the performance on the compound nouns task better predicted word identification accuracy scores. Both of these prosodic sensitivity tasks captured different aspects of reading ability, supporting the idea that different components of prosodic sensitivity, whether due to linguistic level or the specific manipulations of each task, may relate to reading ability differently. These associations provide evidence for isolating and testing components of prosody separately to better measure their effects on reading comprehension.

Researchers have also studied how reading deficits may affect prosodic sensitivity ability. Within the educational system, readers may be classified as reading at or below grade-level, indicating whether reading ability follows a normal pattern of reading skill development for an age group. Researchers Holliman, Sheehy, and Wood (2008) investigated whether poor comprehenders show lower levels of prosodic sensitivity than age and skill matched controls. Prosodic sensitivity was tested in this experiment with four stress related tasks which measured different levels of stress (e.g. syllable, word, phrase, and sentence). The chronological-age matched participants performed better than the poor readers on all prosody measurements, but a main effect of group was found only for the word and syllable level tasks. These results indicate that the ability to effectively detect, use, and understand prosodic characteristics may be a key component in comprehending words and phrases.

These prosodic sensitivity studies, although providing evidence for the relationship between prosodic sensitivity skill and reading comprehension, are limited by their lack of experimental variety. They mainly utilize stress as the prosodic characteristic of choice, which severely limits the wide range of prosodic characteristics utilized in speech. This also diminishes understanding how listeners utilize prosody in various sentence types, which again narrows the component of prosody as its role of a disambiguation tool for comprehension. In our current prosodic sensitivity study, as well as our prosodic fluency task, we ensure that there is a diverse set of stimuli which elicits particular prosodic characteristics for specific types of disambiguation.

The Current Study

Although different studies use different terminology and definitions of prosody, most have generally concluded that there is indeed a relationship between prosody and reading comprehension. We predict the results from our study will also support this relationship, as well as giving new, more specific insight into said relationship. The current study includes an imitation task to measure prosodic fluency and a listening comprehension task to measure prosodic sensitivity. Our constructs from both experiments were informed not only by prior reading comprehension research, but also psycholinguistic research on prosody. Drawing from both backgrounds allowed for an in-depth analysis of prosody's role in reading comprehension, and how specific prosodic features may differ in how they relate to reading comprehension. By having both an imitation task and a listening comprehension task with syntactic and semantic structures set apriori, we can draw better conclusions between prosody and reading comprehension, which could span across both prosodic fluency and sensitivity.

The current prosodic fluency study differed from past fluency and reading comprehension studies in that we utilized an imitation task rather than asking the participants to read sentences aloud. For our purposes, using an imitation task allowed us to stringently control for, and make predictions about, prosodic patterns which may not be present in spontaneous speech (Silverman, & Ratner, 1997). Imitation tasks do not directly measure the process of imitation; without perceptual measures it may be difficult to discern how the participants perceived the model sentences and how those perceptions affected their imitations. What the listeners perceive are the salient characteristics of the sentence will play a role in influencing their imitations (Pardo, 2013). Through the careful creation of our stimuli using past psycholinguistic research, we have attempted to ensure specific prosodic characteristics are present for each construct and that those particular cues are important to the overall prosodic structure. Additionally, a trained ToBI annotator analyzed one model sentence (displayed in Figure 1) example from each construct, to further safeguard that the prosodic phonological structures predicted in the sentences were indeed present. By checking that our prosodic cues were salient in the model sentences and that they were meaningful within the larger sentence structure, we ensured that they may be more likely to perceive and then imitate those characteristics.

Although originally motivated by Miller & Schwanenflugel (2006), our study deviates in the population and the methodology, while still incorporating a similar set of diverse experimental stimuli. In the current imitation study, participants listened to a model sentence and were instructed to produce it. The acoustic data analyzed was measured from the participants' productions. Our stimuli were either ambiguous or unambiguous. Our ambiguous stimuli were created to be disambiguated through prosody, which make those prosodic cues more salient. This experiment contained six distinct constructs, each of which contained a particular syntactic structure. The types of sentences measured were declarative statements, yes-no questions, basic quotatives, ambiguous coordinate structures, relative clauses, and unambiguous coordinates. Within these constructs, F0 change (including F0 variation) and word duration were analyzed.

Our two experiments utilized some overlapping syntactic manipulations and similar stimuli, and importantly, they both used participants from the same population. This allowed us to study both aspects of prosody in reading to better understand how it relates to reading comprehension. A listening comprehension task was used for the prosodic sensitivity experiment: participants listened to an ambiguous stimulus, and were instructed to choose which of the two images shown depicted the sentence. The stimuli were disambiguated by prosodic features, ensuring that accuracy scores reflected successful comprehension from correctly detected prosodic cues.

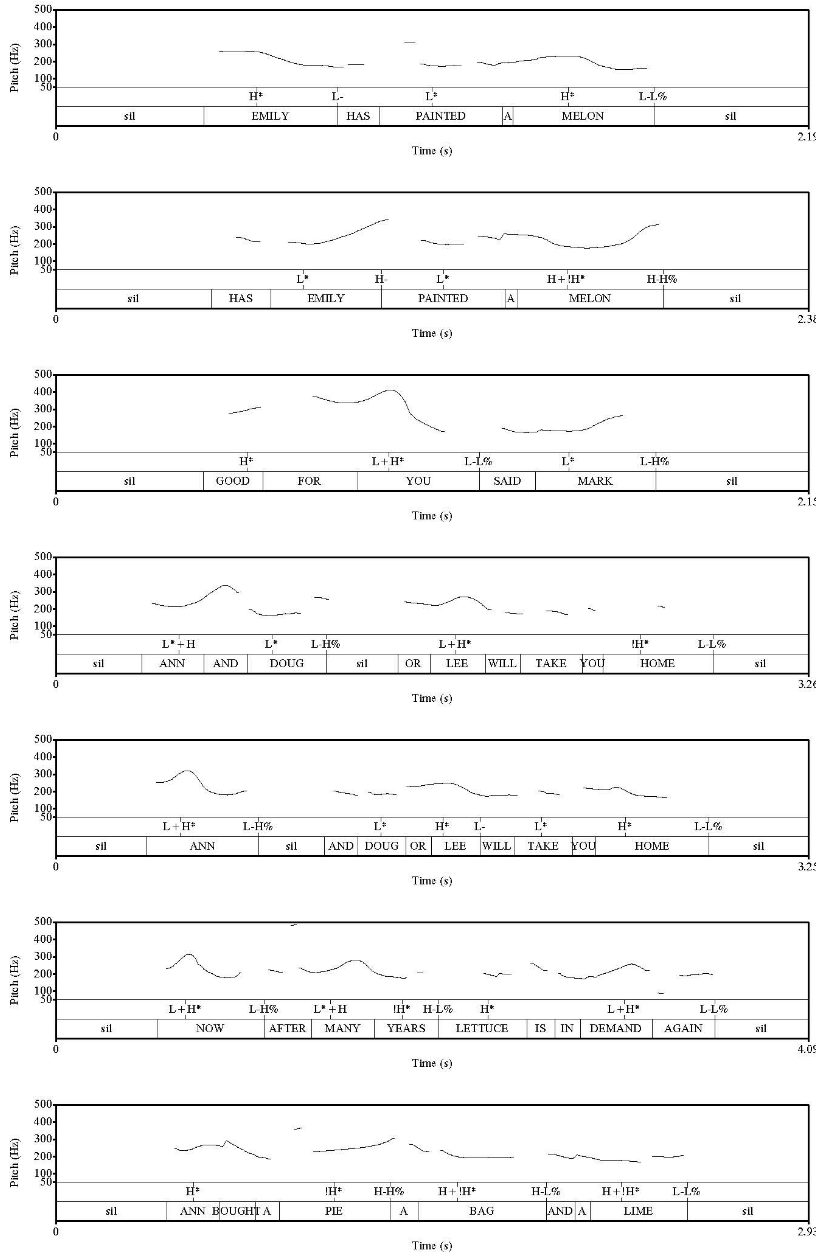

Figure 1.
ToBI annotation (Hz) of one model sentence in each construct: Declarative statement, yes/no question, basic quotative, ambiguous coordinate (2-1), ambiguous coordinate (2-2), relative clause, and unambiguous coordinate structures. From "Prosodic fluency predicts reading comprehension ability," by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

When participants chose the correct image, they must have first attended to and successfully comprehended the sentence, a process which should involve prosodic disambiguation. Seven syntactic constructs were given to participants: Accent disambiguation (AD), focus ambiguity (subject vs. object) (SFOF), contour ambiguity (CA), coordination (CO), phrasing ambiguity (AP), disjunctive questions (DQ), scope ambiguity (SA). However, SA was not included in analysis as the stimuli were not sufficiently ambiguous.

The purpose of our current study is to bring greater clarity to prosody, both prosodic fluency and sensitivity, and its relationship with reading comprehension. From prior research it is clear that there is a relationship between the two, but there is not enough consensus or specificity to attempt a clear unifying conclusion. In part this study will bring about an operationalized definition of prosodic fluency which brings together both the psycholinguistic aspects of prosody as well as the findings from reading comprehension research. The use of specific syntactic stimuli coupled with individual predictions allows for a greater understanding of what prosodic fluency is, and how it is related to reading comprehension. Additionally, the use of high school students gives us insight into how these prosodic features, which are usually developed in elementary grades, impact reading comprehension later. The progression of both reading comprehension skills and prosodic fluency are usually measured in elementary school children and compared with the results from adults, however studying adolescents may provide further information bridging the gap between those two groups. The inclusion of two experiments studying prosodic fluency and prosodic sensitivity expand the analyses we are capable of and allow us to explore the relationship between those measures. We predict that the control comprehenders will be more prosodically fluent than poor comprehenders, in that they will mark the appropriate locations within a sentence with stronger prosodic productions. We also predict that participants with better comprehension ability will be more prosodically sensitive than those with poor comprehension skill, with better accuracy scores in the listening comprehension task.

Experiment 1: Prosodic Fluency

Previous research has shown that reading comprehension skill is related to prosodic fluency: participants with stronger reading comprehension ability also show better prosodic fluency (Benjamin & Schwanenflugel, 2010; Miller & Schwanenflugel, 2006; 2008; Binder et al., 2012). This first experiment was designed to test the hypothesis that poor comprehenders produce prosodic cues less effectively than the control comprehenders. This was done through an imitation task, in which participants were prompted to listen to a model sentence and imitate it. Experiment 1and its participants, materials and results were originally utilized in Breen, Kaswer, Van Dyke, Krivokapić, and Landi (2015).

METHOD

Participants

Twenty-seven participants contributed data to the analyses below. All were high school students between the ages of 13 and 18 at the time of the experiment. These students were recruited by flyers posted through the community and online advertisement, and they were compensated $25 per hour for their participation. None of these students were diagnosed with learning or reading disabilities. All participants were native speakers of American English and were not exposed to any other languages before the age of 7. Only those students who assented to participation and whose guardians gave consent were studied. All procedures were approved by the Yale University Human Investigations Committee. A total of 44 participants completed the experiment, and from that larger set we selected a set of specifically poor comprehenders (N = 12) and a set of control participants (N = 15). The two groups were matched on age, gender composition, and phonetic decoding ability (as measured by the Word Attack subtest (WA) of the Woodcock Johnson III) (Woodcock, McGrew, & Mather, 2001). All participants scored at or above the normal range of performance for IQ (as measured by Weschler Abbreviated Scale of Intelligence II (WASI; Weschler, 1999). The comprehension groups differed in terms of their reading comprehension standard scores, as measured by the Kaufman Test of Educational Achievement II (KTEA; Kaufman & Kaufman, 2004).

We classified participants as specifically poor comprehenders had KTEA standard scores below 95; participants in the control group had KTEA standard scores over 95. Paired sample t-tests were conducted to test for differences between the poor comprehenders and the control group. Table 1 gives the parameters for each group on the measures of age, word attack score, IQ, and KTEA score. Significant differences were found between groups for IQ and reading comprehension skill; no significant differences were found between age and word attack score. Although there is a significant difference between the two groups on IQ, a direct relationship between IQ and language skill has not been found within the normal range of IQ (Nation, Clarke, & Snowling, 2002).

Table 1

Significance Testing between Control and Poor Comprehenders

	Control	**Control**	**Poor**	**Poor**	**t**	**p**
N	15		12			
Number of Females	8		5			
	Mean	**SD**	**Mean**	**SD**		
Age	16.6	1.9	17.6	1.5	$t(25.0) = -1.6$	0.12
Word Attack Standard Score	109	6.7	109	7.1	$t(23.1) = -0.1$	0.92
IQ (WASI)	112	16.6	97.6	2.9	$t(14.0) = 3.07$	<0.01
KTEA Standard Score	110	8.6	84.5	5.0	$t(23.0) = 9.5$	<0.001

Note. From “Prosodic fluency predicts reading comprehension ability,” by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

Materials

In order to test if reading comprehension ability predicted prosodic fluency, we created a set of seven constructs each containing one sentence type. These sentences varied in their syntactic and semantic structure. Each sentence type was designed to elicit the production of a variety of prosodic features including pitch, intensity, and duration. An example of each type of construction appears in Table 2. The sentences were of 6 different types: (a) declarative statements, (b) yes-no questions, (c) basic quotatives, (d) ambiguous coordinate structures, (e) relative clauses, (f) and unambiguous coordinate structures. The full list of experimental stimuli can be found in Appendix A. All constructs contained 10 sentences, except for phrasing which contained 20, and ambiguous coordinate structure, which contained 10 sentences for each of its 2 conditions. The model sentences were all prerecorded by the same trained speaker, whose prosodic productions were controlled to match the predicted prosodic cues.

Table 2

Descriptions and Examples of Construct Types

Construct	Construction Type	Example	Acoustic Features of Interest
1	Declarative Statement	Emily has painted a melon.	F0
2	Yes/No Question	Has Emily painted a melon?	F0
3	Basic Quotative	“That sounds wonderful!” said Jane.	F0
4	Ambiguous Coordinate Structure	Ann(,) and Bobby(,) or Nancy, will come.	Duration
5	Relative Clause	The room, which had a red chair, caught Mandy’s eye.	F0 and Duration
6	Unambiguous Coordinate Structure	Ann has a dog, a pen, and a mug.	F0

Note. From “Prosodic fluency predicts reading comprehension ability,” by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

Declarative Statements and Yes-No Questions

Declarative statements and yes-no questions have contrasting F0 predictions, allowing us to combine the predictions for comparison. Declarative statements are characterized by a final phrase pitch fall, while yes-no questions are realized with phrase final pitch rise (Eady & Cooper, 1986; Pruitt & Roelofsen, 2014). Prior research has shown that large sentence-final pitch declination correlated with stronger reading comprehension skills (Dowhower, 1987; Schwanenflugel et al., 2004). The prior research for yes/no questions has shown the opposite effect, in that a larger final pitch rise is related to better reading skill (Miller & Schwanenflugel, 2006). We predicted that the mean pitch in the final word (*melon*) in a declarative sentence *Emily has painted a melon* would be lower than that of the mean pitch in the final word of *Has Emily painted a melon?*. We also predicted that the larger the pitch change in the final words of each of those sentences (the larger the fall in declaratives and the higher the rise in yes/no questions) would correlate with stronger reading skills. These predictions were extended to predict that the larger difference between the pitch change between declarative and yes-no questions would correlate with better reading skills.

Basic Quotatives

The basic quotative stimuli, such as, "*That sounds wonderful!" said Jane*, included directly reported speech through vocalization verbs such as *said*, *replied*, *responded*. Each stimulus began with the speech, followed by the attributive phrase (i.e. *said Jane*). Jansen, Gregory and Brenier (2001) studied the prosodic features of basic quotatives, which is marked in writing by speech words such as *said*. They found that the directly reported speech (i.e. *That sounds wonderful*) elicits more pitch variation in adult participants than the surrounding attributive phrase. We predicted that the stronger reading skills (skills that were more adult-like) would correlate with a greater difference in pitch variation between the directly reported speech and the attributive phrase; poor comprehenders would have a smaller difference in pitch variation between the speech and the attributive phrase.

Ambiguous Coordinate Structures

The sentences categorized as ambiguous coordinate structure are syntactically complex and globally ambiguous. Within this construct, stimuli were either in one of two conditions: two-one disambiguation by phrasing or two-two disambiguation by phrasing. For example, the sentence (7) was given for both two-one and two-two conditions, but phrased differently. Participants heard both conditions of the stimulus, the two-one disambiguation (7a) and the two-two disambiguation (7b):

(7) Ann(,) and Bobby(,) or Nancy, will come.

(7a) Ann and Bobby, || or Nancy, will come

(7b) Ann, || and Bobby or Nancy, will come

In the two-one disambiguation (7a), the correct interpretation is that Ann and Bobby will come or only Nancy will come. In the two-two disambiguation (7b), the disambiguation of the stimuli indicates that Ann and Bobby will come or Ann and Nancy will come. Previous research has shown that an adult speaker will prosodically group words together in ambiguous series especially with the use of the word “or” (Wagner, 2005; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). These groups require a prosodic boundary, specifically before the conjunct that is not attached to the current phrase. This boundary can manifest prosodically as a lengthening of the word preceding the boundary and following silence. We predict that constituents which precede a boundary will be realized with greater length than those which do not precede a boundary. In the previous example, we predict that *Ann*, the first conjunct, will be longer in the two-two disambiguation (7b) as it precedes a boundary than in two-one disambiguation (7a), whereas *Bobby*, the second conjunct, will be longer in two-one disambiguation (7a) than two-two disambiguation (7b) because it precedes a boundary. Furthermore, we predict that the differences between the boundary and non-boundary nouns between the two disambiguations will be correlated with KTEA scores. We predict poor comprehenders will elicit a smaller difference in durations between disambiguations than the controls.

Relative Clauses

These sentences contained a relative clause, such as the sentence *The room, which had a red chair, caught Mandy's eye*. These clauses (*which had a red chair*) may be interpreted by participants as an informational aside, much in the way parenthetical clauses are treated. Kutik, Cooper and Boyce (1983) demonstrated that the end of the parenthetical clause elicited a pitch change, marking a boundary. Therefore, our predictions focused on the differences in the mean pitch change between phrase-final words (*chair*) and non-phrase final words (*had* and *caught*). Using the precedence that there is a pitch change in the end of a parenthetical clause, the mean pitch change measured for the phrase-final word should be larger than the mean pitch change of non-phrase final words. We predicted that the larger that difference is between the mean pitch changes of the two types of words, the stronger the reading skills.

The relative clause stimuli also included predictions about duration. Kutik et al., (1983) found that the end of a parenthetical clause marked a prosodic boundary through pausing behavior. In light of their findings, we predicted that participants will mark the boundary of the parenthetical clause by lengthening the final word in the clause (*chair*), comparative to non-phrase final words (*had* and *caught*). Although this is not a pause as Kutik et al., (1983) studied, the lengthening of a boundary word and any accompanying silence after has similar prosodic implications to a pause.

Unambiguous Coordinate Structures

The unambiguous coordinate stimuli in the list construct contained lists of nouns such as *Ann has a dog, a pen, and a mug*. Miller and Schwanenflugel (2006) measured the pitch change of the adjectives listed in complex adjectival phrases (*Frog and Toad were happy, playful, curious animal friends*). They found that greater the pitch change within the list of adjectives correlated with stronger reading skills. For our predictions, we measured the average pitch difference between the conjuncts (*a dog* and *a pen*) and non-conjuncts (*Ann* and *has*) of each stimulus. We predicted that better comprehenders also produce a larger difference between the average pitch changes of conjuncts and non-conjuncts than poor comprehenders.

Duration is another prosodic feature which can indicate the presence of a boundary. Previous research has shown that lengthening the constituent preceding a boundary and pausing afterwards can indicate a series of nouns (Dankovicová et al., 2004). In our study, duration was measured by calculating the average duration and following silence (seconds) of the conjuncts within the list (*a dog* and *a pen*) and the non-conjuncts (*Ann* and *has*). We predict that participants will produce greater durations for the conjuncts than the non-conjuncts, however, we also predict that better comprehenders will produce a larger durational difference between the conjuncts than the non-conjuncts compared to the poor comprehenders.

Testing Materials

Prior to testing, participants provided information about their educational, familial, medical, and language history. Sensory testing was conducted with each participant prior to behavioral testing to ensure that hearing and vision were within normal limits. Visual acuity was tested using the Snellen Eye Chart for every participant who was not already wearing corrective lenses. Participants' hearing was tested using a standard audiometer, with all participants accurately identifying 5, 1, 2, and 4 kHz at 20dB.

The Word Attack component of the Woodcock Johnson was used to measure the participants' ability to apply single-word decoding and phonetic skills to unfamiliar nonsense words (Woodcock et al., 2001). Participants were asked to read out loud non-words presented in lists of 5 or 6, each list increasing in complexity. The scores used for analysis were individual standard scores.

Reading comprehension was measured through the KTEA. This subsection of the KTEA is composed of passages of increased difficulty, in which students would answer comprehension questions about the previous passage. Standard scores of this assessment were used for analysis. This test has garnered high levels of internal consistency for its comprehension subsets (such as reading comprehension) with coefficients ranging from .92 to .95 (Worthington, 1987). Additionally the KTEA was shown to be correlated with several other widely used achievement tests (e.g. Wide Range Achievement Test, Peabody Individual Achievement Test, Stanford Achievement Test etc.) (Worthington, 1987).

Intelligent quotient was measured using the Weschler Abbreviated Scale of Intelligence (WASI_P_IQ) (Weschler, 1999). This standardized test measures fluid reasoning, spatial processing, attention to detail, and visual-motor integration. The WASI-II is composed of two subtests: Block Design and Matrix Reasoning. The Block Design tests visual perception and spatial reasoning through assessing the ability to analyze abstract figures and create a pattern from a model (WASI-II Manual). Participants view a model and must recreate the design using blocks provided to them. The Matrix measures nonverbal fluid reasoning, inductive reasoning, and spatial reasoning by testing participants' abilities to mentally manipulate abstract symbols and understand the relationships among them. Participants view an incomplete matrix or series of symbols and are asked to complete the series. There are four types of items on this subtest: pattern completion, classification, analogy, and serial reasoning (WASI-II Manual).

Procedure

Before experimental testing began, the participants completed the demographic questionnaire as well as three measures of standardized testing. These included the Word Attack subset of the Woodcock Johnson, the WASI, and the KTEA. This preliminary testing took an average of 40 minutes.

After gaining the participants' assent, they were seated before a computer screen which visually presented the stimulus sentences for the participant while imitating the sentences, so the participants did not have to actively remember the sentences. Participants heard the stimuli through headphones provided by the researchers. Productions were recorded using a Sennheiser ME66 cardioid microphone positioned beside the monitor, turned in the direction of the participant's head. Sound tests were performed prior to experimentation using Praat speech analysis software to verify mic functionality and a frequency sampling rate above 44,100Hz. Directions and stimuli were displayed using e-Prime version 2.0.

The experiment began with instructions explaining the concept of prosody as well as the production expectations of the participant. They started with an audio recording of a male voice saying the sentence "Mary came home" with the appropriate prosody of a declarative statement. Participants were instructed to listen to the stimulus sentence and were then given both an appropriate and an inappropriate example of reproduction. The first example exemplified a "good participant” as it paralleled the original sentence in content and melody. The second example was categorized as a “bad participant” as it was produced with the prosody of a question (i.e. a sentence final pitch rise).

After the example demonstration concluded, participants were given time to ask questions. Participants were then given directions for how to record their responses and how to control the presentation of the stimuli. Instructions were presented on the screen and prompted the participant to listen to the recorded female voice through the headphones and to repeat the sentence aloud into the microphone, while highlighting the importance of reproducing the prosodic structure. Each stimulus could be played a maximum of twice before recording. After listening to the stimulus, participants were prompted to press a key to record their response. Once the participant finished producing the entire stimulus they pressed a key to stop recording, and then pressed that same key again to continue onto the next stimulus. The participants were monitored by the researcher in an adjacent testing room during the experiment.

The experiment contained three blocks of a total of 270 stimuli with 90 experimental stimuli. The same 90 experimental stimulus sentences were presented randomly within each of the three blocks. The experiment took approximately 45 minutes to complete. A total of 2 breaks were given to participants between each block of stimuli; participants were encouraged to take the breaks but could choose to opt out.

RESULTS

Each of the 90 sentences presented to participants included three versions which were produced by 28 speakers, resulting in a possible 7560 production total. Five files (one from one participant and four from another) were not included because the production was unidentifiable, leaving 7550 productions for analysis. We used the Prosodylab-Aligner (Gorman, et al, 2011) to force align the words from the target stimuli with their waveforms. Using Praat (Boersma & Weenink, 2011), we extracted a set of acoustic features from each word in each sentence. These features included pitch (mean F0, maximum F0, and F0 values at 10 equal-spaced intervals across the word), and duration (defined as the duration of the word itself and any following silence).

A z-score transformation was utilized to standardize the individual scores to measure across acoustic features. For every feature (pitch and duration) from every individual participant, the value associated with the feature was subtracted from each participant's individual average for that particular feature, and was then divided by that participant's feature standard deviation. Any values which were more than 3 standard deviations were excluded from analysis.

The goals of this experiment include quantifying prosodic fluency in order to create a more cohesive definition, as well as exploring the relationship between prosodic fluency and reading comprehension. The acoustic features taken from the imitation productions were analyzed for the presence of prosodic characteristics that have been previously studied in respective sentence structures, providing tools for quantification. Additionally, these prosodic characteristics have been noted to significantly differ based on reading comprehension ability, either detected through testing or by age group (i.e. adult compared to children). Our statistical analyses were conducted to test whether participants were producing the prosodic features associated with each construct, and whether the poor comprehenders would realize those features significantly less effectively than the control group.

We used mixed-effects linear regression models with trial-by-trial data to test our predictions. The dependent variable in each model was the continuous acoustic measure, and the fixed effects were the sentence type (the construct), the comprehension group (poor or control), and the interaction of those two factors. Random effects of participant and item were included in each model. We attempted to fit a fully saturated model with random intercepts for subject and item as well as random slopes for both the main effects and the interaction, however this model never converged. We tested for the model of best fit by comparing each model to a less complex model found within it to determine whether the additional terms in the random effects structure were justified according to the procedure specified by Baayen (2008). The more saturated model was only used if it significantly improved model fit. A full table of the parameter estimates of each model can be found in Table 3.

Table 3

Parameter Estimates of Mixed Effects Models for Experiment 1

Syntactic Contrast: Acoustic Feature		Est.	SE	t	p
Statement vs. Yes-No Q: Pitch					
	Intercept	-0.45	0.11	-4.11	*
	Statement vs. Yes-No	-1.27	0.16	-7.75	*
	Group (Control and Low)	-0.15	0.12	-1.22	n.s.
	Sentence Type x Group	-0.06	0.24	-0.24	n.s.
Basic Quotative: Pitch					
	Intercept	0.22	0.06	3.57	*
	Attributive (non-quote) vs. Quote	-0.33	0.09	3.64	*
	Group (Control and Low)	-0.03	0.05	-1.45	n.s.
	Word category x Group	-0.12	0.08	1.52	n.s.
Two-Two and Two-One Phrasing: Duration 1st constituent					
	Intercept	0.41	0.10	3.95	*
	Boundary vs. No boundary	1.06	0.20	-5.36	*
	Group (Control and Low)	0.14	0.09	1.64	n.s.
	Boundary x Group	0.30	0.20	-1.53	n.s.
Two-Two and Two-One Phrasing: Duration 3rd constituent					
	Intercept	0.55	0.15	3.75	*
	Boundary vs. No boundary	0.76	0.10	7.42	*
	Group (Control and Low)	0.14	0.07	2.15	*
	Boundary x Group	0.46	0.07	6.93	*
Relative Clause: Pitch					
	Intercept	0.01	0.05	0.21	n.s.
	Boundary vs No Boundary	0.69	0.05	13.42	*

	Group (Control and Low)	0.07	0.04	1.76	n.s.
	Boundary x Group	0.19	0.10	1.86	^
Relative Clause: Duration					
	Intercept	0.49	0.07	6.99	*
	Boundary vs No Boundary	1.72	0.14	11.93	*
	Group (Control and Low)	0.05	0.04	1.13	n.s.
	Boundary x Group	0.28	0.10	2.69	*
Unambiguous Coordinate: Duration					
	Intercept	0.35	0.07	4.77	*
	Boundary vs. No Boundary	-1.08	0.08	-12.77	*
	Group (Control and Low)	0.00	0.03	-0.07	n.s.
	Boundary x Group	-0.20	0.06	-3.27	*
Unambiguous Coordinate: Pitch					
	Intercept	0.19	0.04	4.36	*
	Conjunct vs Non-conjunct	-0.48	0.05	-9.15	*
	Group (Control and Low)	-0.08	0.05	-1.58	n.s.
	Word category x Group	-0.10	0.11	-0.98	n.s.

Note. Random slopes were used in each model, so p-values cannot be calculated; * indicates estimated significance beyond the .05 level; ^ indicated marginal significance; n.s. indicates a non-significant effect. From "Prosodic fluency predicts reading comprehension ability," by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

Declarative Statements and Yes-No Questions: F0

The constructs of declarative statements and yes-no questions are characterized by contrasting sentence final F0 contours; combining their analyses proved useful in strengthening our findings. We did this by comparing the normalized average F0 slopes of the final word in each stimulus in both sentence types across groups. The mean F0 slope of the final word in both sentence types was calculated by subtracting the F0 measured at 80% of the word produced from the F0 measured at 20%. A positive F0 slope indicated a F0 fall while a negative F0 slope signaled a F0 rise.

Although all participants were predicted to produce sentence final pitch falls for declarative statements and sentence final F0 rises for yes-no questions, we predicted that the size of the difference between the two F0 slopes would be modulated by the KTEA groups. More specifically, that the control group would produce a significantly larger difference between the F0 slopes of the two conditions as compared to the poor comprehenders. The results of the mixed-effects linear regression with F0 slope as the dependent variable revealed a main effect of the construct manipulation (statement vs. yes-no question), indicating that construct significantly predicted the F0 slope ($t = -7.75*$). Statement stimuli produced more positive F0 slopes while yes-no questions produced more negative F0 slopes. There was no main effect of group type; both the controls and the poor comprehenders realized comparable F0 slope differences ($t = -1.22$). There was also no interaction between group type and stimulus type ($t = -0.24$). The pitch slope of model speaker, and the poor and control comprehenders is displayed in Figure 2.

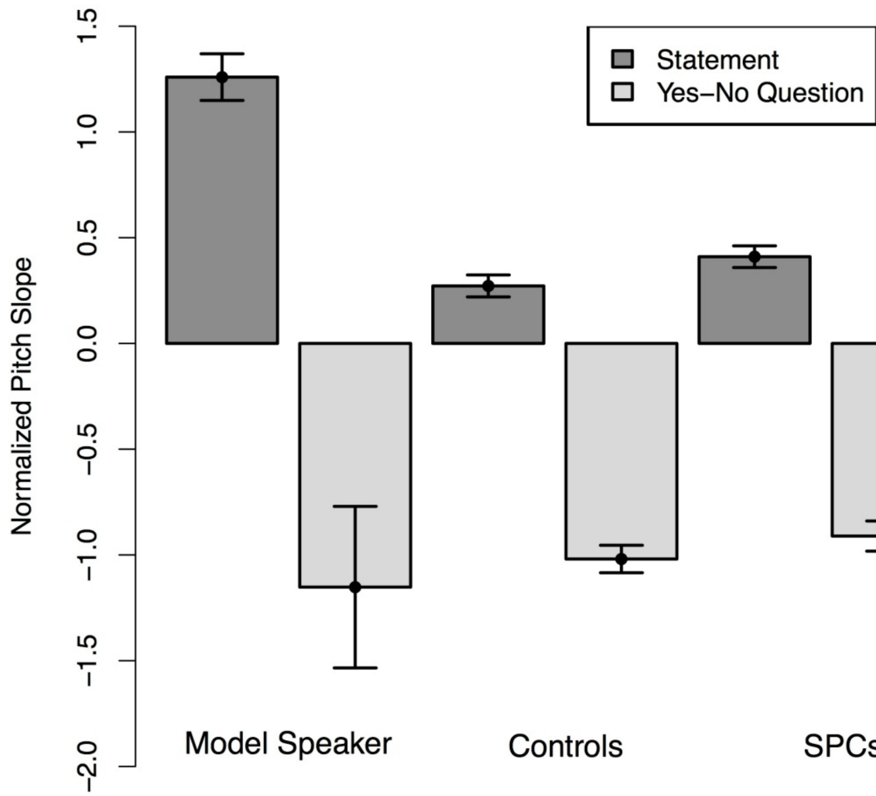

Figure 2. Normalized pitch slope of the final word of Statements and Yes-No Questions as produced by the Model Speaker, Control comprehenders, and SPCs (poor comprehenders). A positive pitch slope corresponds to a declination in F0 over the word; a negative slope corresponds to a F0 rise. Error bars represent standard errors. From "Prosodic fluency predicts reading comprehension ability," by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

Basic Quotatives: F0

The basic quotative type utilized F0 as the dependent prosodic measure. This construct contains differences in F0 variability, with directly reported speech eliciting larger variations in F0 than the attributive phrase. We assessed the F0 variability differences between the two components by averaging the difference between the minimum and maximum F0's of each word within the quote (e.g., *That*, *sounds*, and *wonderful*), and again averaging the difference between the minimum and maximum F0's of the F0 variation of the words of the attributive phrase (e.g. *said* and *Jane*). We predicted that the difference between the two portions of the basic quotative stimuli would be reflected in the comprehension group type. The participants in the control group would produce greater F0 variation in the directly reported speech than in the attributive phrase as compared to the difference in F0 variation between the two components for the poor comprehenders.

The mixed-effects linear regression included phrase type (directly reported speech or the attributive phrase), comprehension group type, and the interaction of the two factors on the F0 variability measure. The analysis showed a significant main effect of phrase type on F0 variation ($t = 3.64$*), with the quote group eliciting a significantly larger F0 variation than the non-quote group (i.e the attributive phrase). There was no main effect for comprehension group on F0 variation ($t = -1.45$), but there was a suggestion of an interaction between the two factors ($t = 1.52$). The quotes elicited a larger F0 variability value than the attributive phrases while the poor comprehenders did not have as large of a difference. Figure 3 shows the F0 variation between groups for the basic quotative construct between comprehension groups and model speaker.

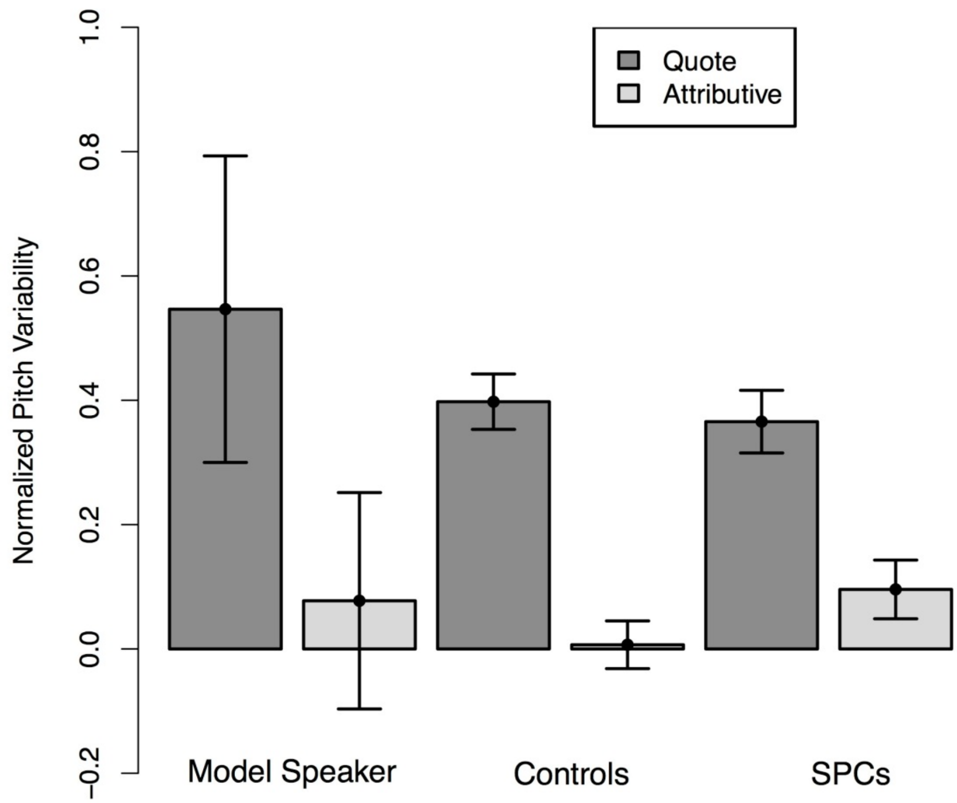

Figure 3. Normalized pitch variability on the quote and attributive phrases of Simple Quotative sentences as produced by the Model Speaker, Control Speakers, and SPCs (poor comprehenders). Error bars represent standard errors. From "Prosodic fluency predicts reading comprehension ability," by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

Ambiguous Coordinate Structures: Duration

This construct contained two possible interpretations for each stimulus (e.g. *Ann(,) and Bobby(,) or Nancy, will come*) based on the placement of the prosodic boundary: two-two disambiguation and two-one disambiguation. Duration plus any following silence was used to indicate the presence of a boundary; specifically the noun which preceded the predicted prosodic boundary was measured. This measure was normalized for the analyzed constituents. The two-one disambiguation produced *Ann and Bobby* || *or Nancy will come* (Ann and Bobby will come or only Nancy will come) and the two-two disambiguation produced *Ann* || *and Bobby or Nancy will come* (Ann and Bobby will come or Ann and Nancy will come). We measured the durational differences of *Ann* across both interpretations, as well as *Bobby*, as both words precede a boundary depending on the prosodic contour.

We predicted overall that the constituents of interest in the same position as *Ann* (the first conjunct) will be longer in two-two disambiguation than in two-one disambiguation, whereas those constituents with position placements like *Bobby* (the second conjunct) will be longer in two-one disambiguation than two-two disambiguation. Furthermore, we predict that the differences in the first conjunct between the two disambiguations will be affected by comprehension group type: the control group will produce a larger difference in duration between the two disambiguations than the poor comprehender group. Again, we hold that same prediction for the second conjunct.

Mixed-linear effects models were run for condition (two-one or two-two) and comprehension group type on duration for the first conjuncts (*Ann*) and the interaction between the two factors. There was a main effect of condition on duration length (t = -5.36*), with participants lengthening the first conjunct in the two-two condition significantly more than in the two-one condition, as predicted there was a marginal main effect of group type on duration (t = 1.64), as the control group produced the constituents with greater duration across conditions compared to the poor comprehenders. Additionally, a suggestive interaction was found between the factors (t = -1.53) due to the control group producing a larger duration difference between constituents than the poor comprehender group. Figure 4 displays the mean normalized duration of the model speaker, and the poor and control comprehenders for the first conjunct for each condition of the ambiguous coordinate structures.

The same model was run for the second conjunct (*Bobby*) which included condition type (two-one or two-two) and group type on duration and the interaction. There was again a main effect of condition on duration length (t = 7.42*), with participants lengthening the second conjunct in the two-one condition significantly more than in the two-two condition. There was a main effect of group (t = 2.15*), with the control group producing the second conjunct with a significantly longer duration across both conditions as compared to the poor comprehenders. A significant interaction was found between condition type and group type (t = 6.93*), which resulted from the controls producing a larger durational difference between the conditions than the poor comprehenders. Figure

5 displays the mean normalized duration of the model speaker, and the poor and control comprehenders for the second conjunct for each condition of the ambiguous coordinate structures.

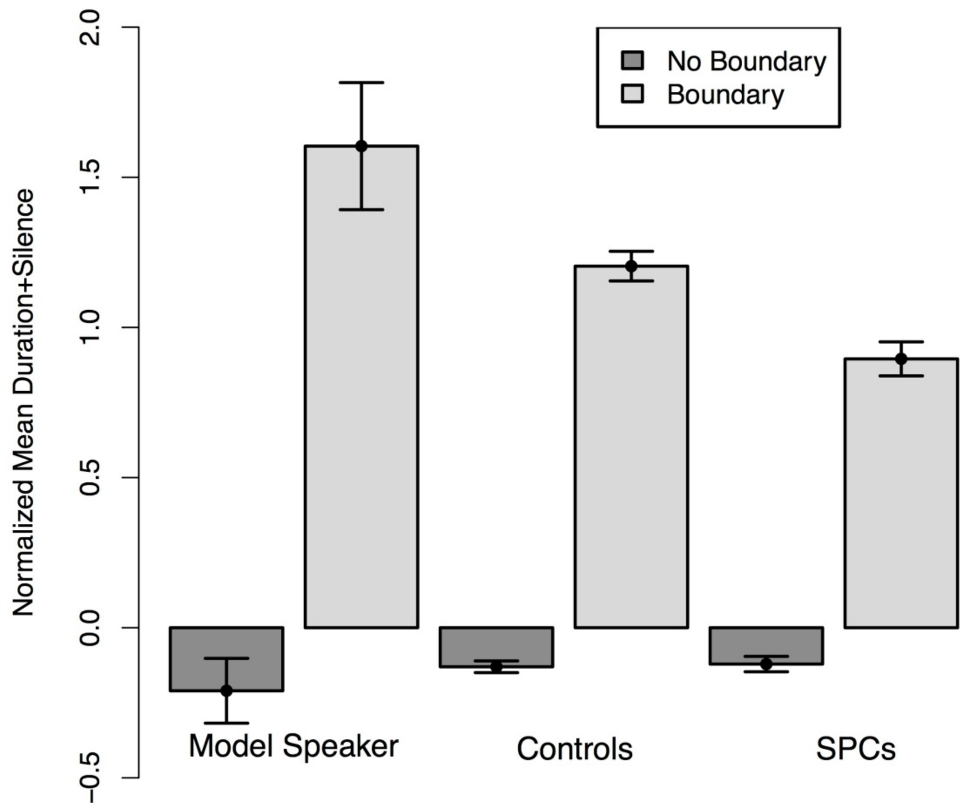

Figure 4. Normalized average duration plus silence of the first conjunct (e.g. *Ann*) of the Ambiguous Coordinate Structure sentences depending on whether the speaker intended the two-one (No Boundary) or two-two (Boundary) structure as produced by the Model Speaker, Controls, and SPCs (poor comprehenders). Error bars represent standard errors. From "Prosodic fluency predicts reading comprehension ability," by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

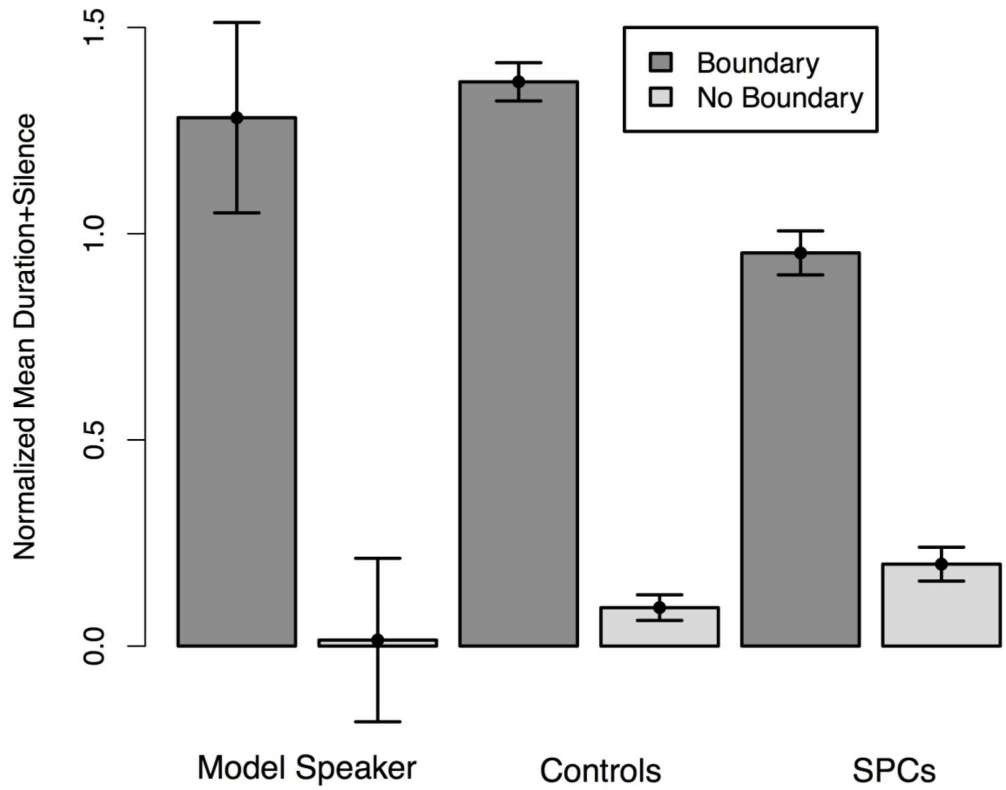

Figure 5. Normalized average duration plus silence of the second conjunct (e.g. *Bobby*) of the Ambiguous Coordinate Structure sentences depending on whether the speaker intended the two-one (Boundary) or two-two (No Boundary) structure as produced by the Model Speaker, Controls, and SPCs (poor comprehenders). Error bars represent standard errors. From "Prosodic fluency predicts reading comprehension ability," by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

Relative Clauses: F0

The relative clause construct (e.g. *The room, which had a red chair, caught Mandy's eye*) used F0 as its dependent acoustic feature, in how it was used to signal a prosodic boundary. The realization of appropriate boundaries was analyzed through the differences in the mean F0 change between phrase-final words (*chair* and *room*) and non-phrase final words (*had* and *caught*). The phrase final words were predicted to be adjacent a boundary which can be prosodically manifested in a F0 change. Non-phrase final words that do not precede a boundary and are less likely to elicit a F0 change. Mean F0 change for phrase-final constituents (*chair*) as well as non-phrase final conjuncts (*had* and *caught*) were calculated by finding the difference between the minimum and maximum F0 and normalized the difference before averaging it for each participant. Each participant had a normalized average for the F0 change for phrase-final or non-phrase final words.

For all participants, we predicted that words which were adjacent to a boundary would elicit a larger F0 difference than words which were not adjacent to boundaries. Additionally, we predicted that the control group would produce a larger difference in F0 change dependent on boundary location than the poor comprehenders.

Our analyses, with mean F0 change as the dependent variable, showed a main effect of boundary placement (adjacent or non-adjacent) on F0 change (t = 13.4*), with conjuncts adjacent to boundaries eliciting a larger F0 change than non-adjacent words. A marginal effect of comprehension group was observed (t =

1.76), as controls produced words across both boundary conditions with higher F0 than the poor comprehenders. Figure 6 illustrates the mean normalized duration and silence of the model speaker, and poor and control comprehenders for each location in the relative clause construct.

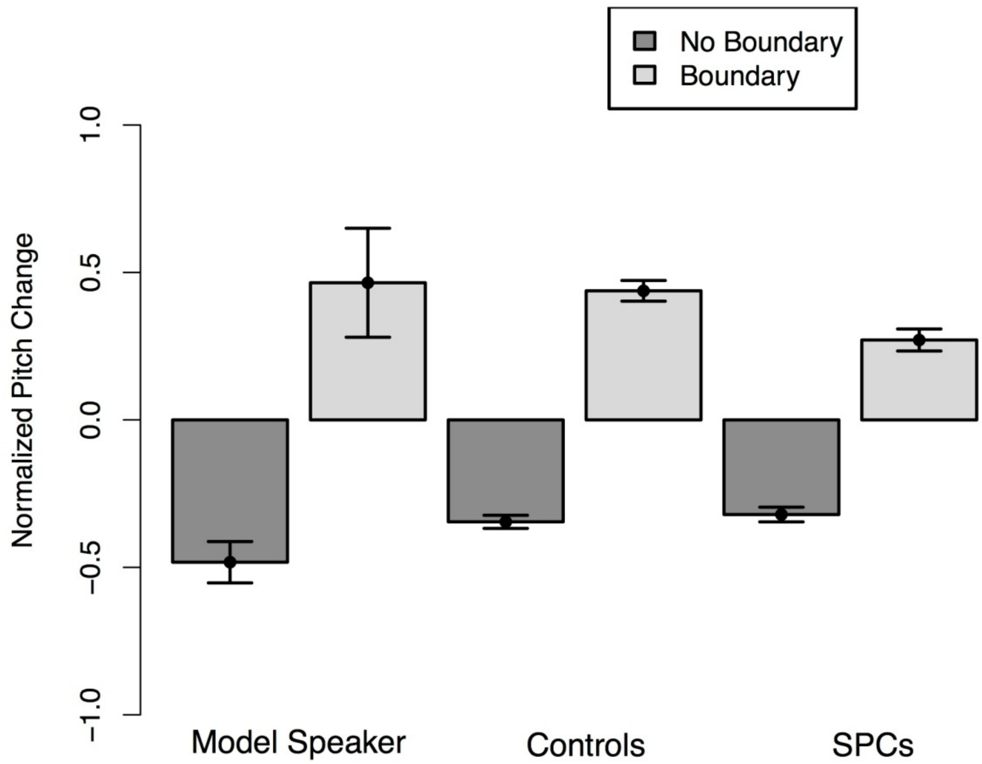

Figure 6. Normalized pitch variation of words preceding hypothesized boundary locations and non-boundary locations in the relative clause sentences as produced by the Model Speaker, Controls, and SPCs (poor comprehenders). Error bars stand for standard error. From "Prosodic fluency predicts reading comprehension ability," by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

A suggestive interaction between group and boundary location (t = 1.86) indicated that controls produced a larger F0 change difference, with the adjacent words eliciting a greater F0 change than non-adjacent, as compared to the poor group.

Relative Clauses: Duration

The prosodic realization of boundaries can include more than one acoustic feature. In the case of the phrasing construct, duration was also utilized to indicate the production of prosodic boundaries. Again, we analyzed the computed average normalized duration and following silence for boundary adjacent words and non-adjacent boundary words in stimuli for each participant.

We predicted that every speaker would realize words that were boundary adjacent with longer duration than those which were not boundary adjacent. We also predicted that participants in the control group would produce larger differences in the duration of boundary adjacent and boundary non-adjacent constructs than the poor comprehenders.

Duration was predicted using boundary location (boundary adjacent, non-adjacent) and group in a mixed effects linear regression model which resulted in a main effect of boundary (t = 11.93*) with participants overall producing words which preceded a boundary with significantly longer duration than those which did not precede a boundary. No main effect of group was observed (t = 1.13), however a significant interaction was found (t = 2.69*). The interaction was due to the controls producing significantly larger durational differences between words at the two boundary locations as compared to the poor comprehenders. Figure 7 depicts the mean normalized pitch variation for the model speaker, and the poor and control comprehenders for each location of the relative clause construct.

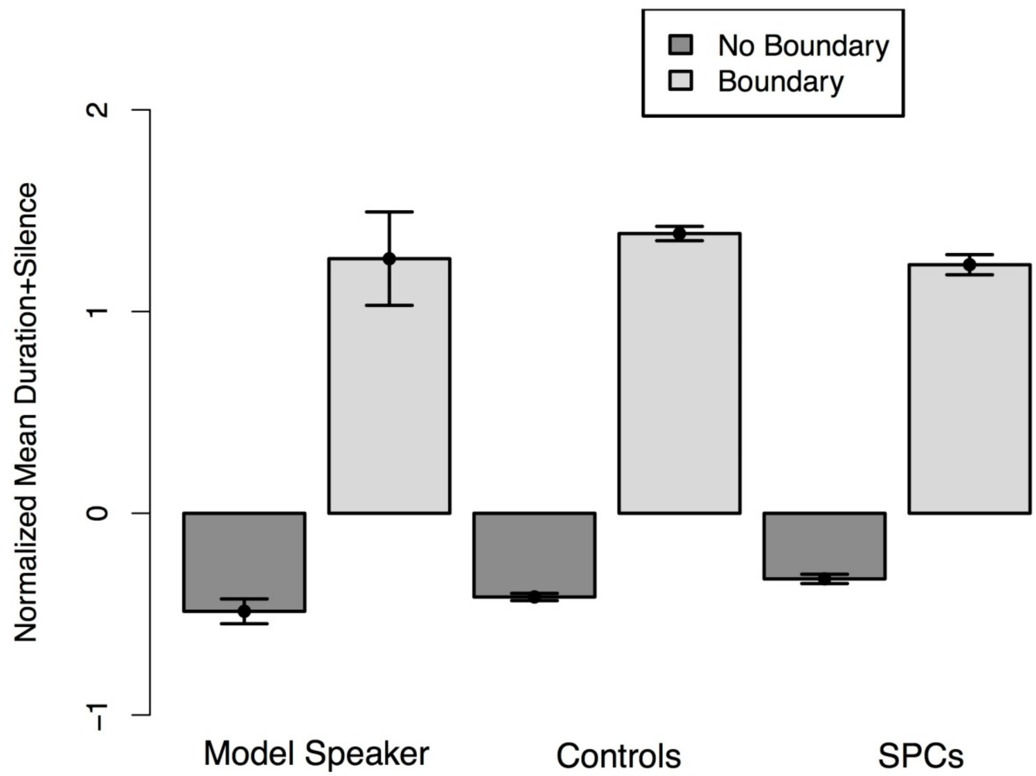

Figure 7. Normalized average duration plus silence of words preceding hypothesized boundary locations and non-boundary locations in the relative clause sentences as produced by the Model Speaker, Controls, and SPCs (poor comprehenders). Error bars represent standard error. From “Prosodic fluency predicts reading comprehension ability,” by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

Unambiguous Coordinate Structures: F0

Prosodic differences were also analyzed in the unambiguous coordinate construct through the feature of F0 change. As characterized earlier, this condition contains conjuncts, the words within a list series (*dog* and *pen*) and the non-conjuncts outside of the list (*Ann* and *has*). Boundaries can be realized through different prosodic characteristics, and both F0 change as well as duration were features we wanted to explore. F0 change was calculated by finding the difference between the minimum F0 and maximum F0 for each of the conjuncts and non-conjuncts. Those differences were normalized and averaged into one F0 change value per participant per boundary type. We predicted the conjuncts, as they preceded boundaries in the list, would produce a greater F0 change than non-conjuncts.

For all participants, we predicted that the production of conjuncts would have greater F0 change than non-conjuncts. We also predicted that comprehension group would modulate those differences, with the controls producing a greater F0 change difference between word types than the poor comprehenders. F0 change was predicted using a mixed-effects linear regression with boundary location (boundary or no-boundary) and group type. A main effect of boundary location was observed ($t = -9.15$*), with all participants producing the conjuncts (boundary) with greater F0 change than the non-conjuncts. There was no effect of group ($t = -1.58$) nor was there an interaction between boundary location and group ($t = -0.98$). Figure 8 displays the normalized mean pitch variability in the model speakers and for both the poor and control comprehenders for each location of the unambiguous coordinate structure.

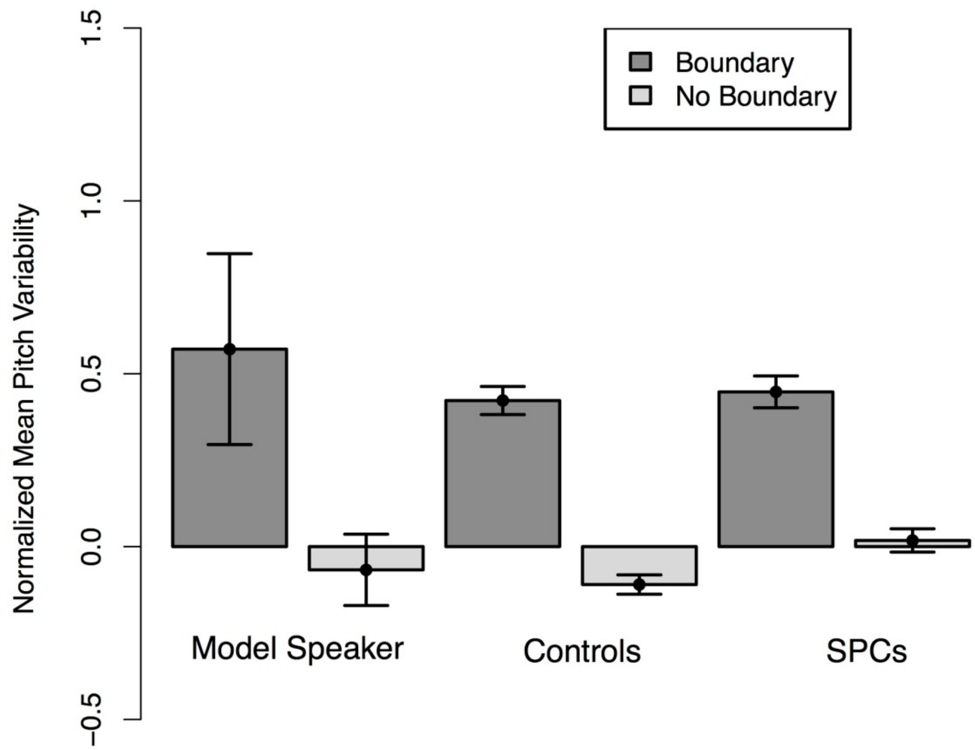

Figure 8. Normalized pitch variation of words preceding hypothesized boundary locations and non-boundary locations in the unambiguous coordinate sentences as produced by the Model Speaker, Controls, and SPCs (poor comprehenders). From "Prosodic fluency predicts reading comprehension ability," by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

Unambiguous Coordinate Structures: Duration

Duration was another prosodic characteristic used to test for prosodic fluency differences between reading comprehension groups in the unambiguous coordinate construct. We predicted that the conjuncts (boundary) would show have longer durations than non-conjuncts (no-boundary). We also predicted that better comprehenders would have greater durational differences between the conjuncts and non-conjuncts than the poor comprehenders. Duration was predicted using a mixed-effects linear regression with boundary location (boundary or no-boundary) and group type. A significant main effect of boundary location was observed ($t = -12.77$*), with all participants producing the conjuncts (boundary) with greater F0 change than the non-conjuncts. There was no effect of group ($t = -0.07$), however, a significant interaction was found between boundary location and group ($t = -3.27$*). The control reading comprehenders produced larger durational differences between boundary and no-boundary groups than the poor comprehenders. Figure 9 displays the normalized duration and silence in the model speakers and for both the poor and control comprehenders for each location of the unambiguous coordinate structure.

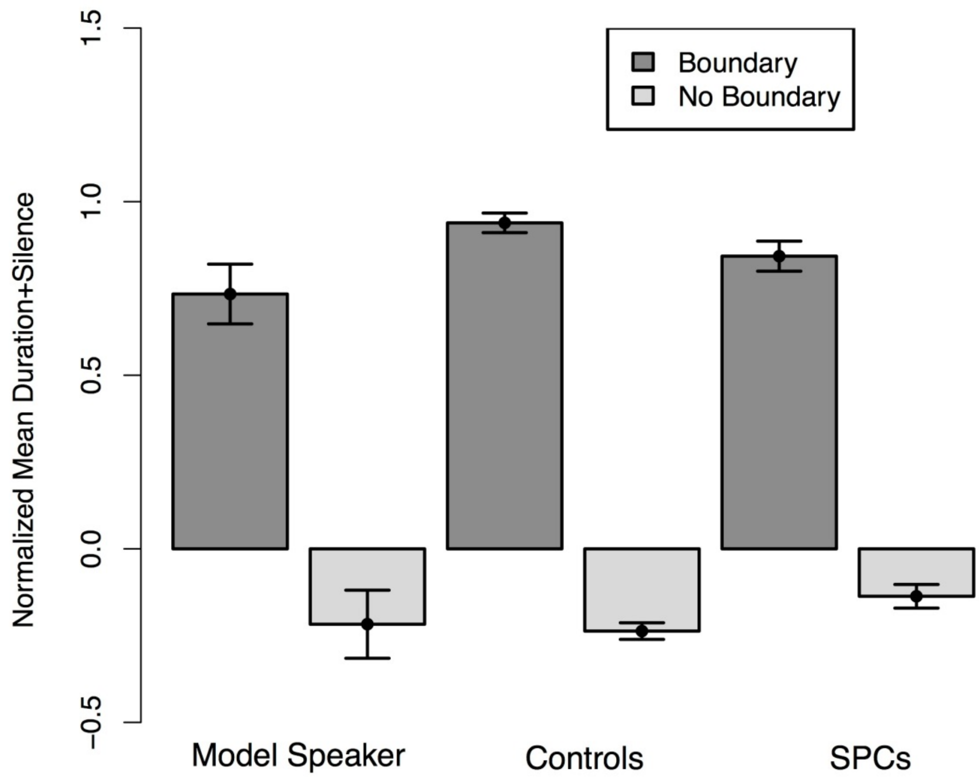

Figure 9. Normalized average duration plus silence of words preceding hypothesized boundary locations and non-boundary locations in the unambiguous coordinate sentences as produced by the Model Speaker, Controls, and SPCs (poor comprehenders). Error bars stand for standard error. From “Prosodic fluency predicts reading comprehension ability,” by Breen, Kaswer, Van Dyke, Krivokapić, and Landi, 2015, *Manuscript submitted for publication*. Reprinted with permission.

DISCUSSION

Of the six syntactic constructs tested in the production experiment, only one proved to elicit significant results for both factors and the interaction factor. The ambiguous coordinate structure, for the analysis of second conjunct (*Bobby*), revealed a significant main effect of condition, reading comprehension group, and a significant interaction. Recall that this structure had two conditions for the example stimulus (11).

(11) Ann(,) and Bobby(,) or Nancy, will come.

In the 2-1 condition, a prosodic phrase boundary was placed after *Bobby*, resulting in the expected disambiguation of the sentence as *Ann and Bobby* will come or only *Nancy* will come. The 2-2 condition, with a boundary after *Ann*, should disambiguate the sentence as either *Ann and Bobby* will come or *Ann and Nancy* will come.

Our analyses focused on only one prosodic characteristic, duration, as the prosodic cue used to disambiguate between the two different conditions in this construct. We tested different prosodic characteristics between our six constructs, but the singling out of this one characteristic in our results lends support that specific features of prosody should be tested and measured individually. These results may indicate that duration has a relationship with reading comprehension which differs from the relationship between reading comprehension and F0. The previous psycholinguistic literature on prosodic boundaries has shown that they facilitate sentence comprehension (Kjelgaard & Speer, 1999; Snedeker &

Trueswell, 2003). However, the literature focusing on prosody, specifically in terms of pausing, and reading comprehension prove to be inconsistent with our results, but this may be explained through differences in participant population and methodology.

The results from our acoustic measures of the ambiguous coordinate structure suggest that the intended syntactic manipulation produced its desired effect in that we found differences in the predicted prosodic characteristic. This helps support any further conclusions drawn, as we know there is an inherent difference between the two conditions based in the placement of prosodic phrase boundaries set from prior psycholinguistic literature (Wagner, 2005).The differences found between the reading comprehension groups also highlight baseline differences important to the overall findings; in other words, participants at control level comprehension ability produced stronger prosodic boundaries than those identified as poor comprehenders. Therefore our results supported our prediction that better comprehenders will better emulate adult-like (i.e. skilled) prosody.

The significant interaction between the two factors of comprehension group and disambiguation condition is the crux of the finding, as it directly provides support for not just the production of prosodic cues, but for the appropriate placement. The control group elicited a greater durational difference between conditions in the predicted pattern; their duration of second conjunct in the 2-1 disambiguation was significantly longer than in the 2-2. The poor comprehenders did not produce such a difference between the two conditions, although they did produce second conjunct in the 2-1 disambiguation longer than the 2-2. Poor comprehenders may not have produced the same strength of duration difference because they did not fully understand the semantic ambiguity and how the prosody alleviated it. This possible conclusion rests on the differences between perceptions of the stimuli by participants. The literature on the efficacy of imitation tasks postulates that the features of the model sentence which are perceived as salient are more likely to be successfully imitated (Pardo, 2013). Despite those cues being salient in our model sentences, differences still arose between imitations of poor and control comprehenders, which may possibly point to the differences in comprehension ability as a cause.

This understanding of the semantic ambiguity is crucial to appropriate boundary placement, as demonstrated in the psycholinguistic literature. Both semantic and syntactic understanding can drive boundary placement, especially in globally ambiguous environments (Breen, et al., 2011; Snedeker & Trueswell, 2003). The poor comprehenders in our study may have attempted to mimic the model speaker, but without the full comprehension of the disambiguation which the prosodic phrase boundary provides, those participants produce a weaker boundary. This is contrasted with the control comprehenders, who not only mimic the model speaker, but possibly due to their better comprehension abilities, recognize the ambiguous nature of the sentence and consequently emphasize the prosodic boundary, as it disambiguates the grouping. A similar effect was seen in Dankovicová (2004), as the elementary school participants produced prosodic boundaries, however they were not as strong as those produced by adults, possibly indicating the younger participants had some understanding of the disambiguating role of the prosodic boundary but it was not as advanced as the adults' understanding.

The thread of duration as the prosodic characteristic of measurement for studies on boundary placement is not surprising given our results, especially that the only other constructs which elicited significant interactions also measured durational differences (unambiguous coordinate construct and relative clause construct). Duration and pausing are also well recorded prosodic characteristics of prosodic phrase boundaries (Breen, et al., 2011; Dankovicová, et al., 2004; Snedeker & Trueswell, 2003; Snow, 1983). Additionally, the isolation of duration as the only prosodic characteristic which attained significant effects and an interaction may indicate that prosodic characteristics should not be treated under one umbrella. In other words, different prosodic characteristics may have different relationships with reading comprehension ability.

Two studies which parallel some this current study's elements, Miller and Schwanenflugel (2006) and Binder et al., (2013), found somewhat conflicting results. The prosody and reading comprehension study by Miller and Schwanenflugel (2006) gave elementary school participants (third graders approximately 8 years old) a passage with a rich variety of sentences: declaratives, basic quotatives, wh-questions, yes-no questions, complex adjectival phrase commas, and phrase-final commas. The researchers examined how reading skills interacted with prosodic production, and they found that the prosodic characteristics of pausing and pitch were associated with reading skill.

Specifically, better readers produced more adult-like F0 change for declarative statements and yes-no questions, and they produced shorter and fewer pauses. In our study, we did not find F0 change between declarative statements and yes-no questions to produce a difference between comprehension groups, nor did we find that more skilled readers made shorter pauses.

The ambiguity present in the ambiguous coordinate structures may have played a key role in the prosodic cues produced. Those sentences were created to ensure that they could be disambiguated through prosodic cues, and so the prosody present in the model sentence guided the participants' disambiguations. Therefore, the participants who were more skilled at comprehending the role which the prosodic boundaries played in disambiguation could be more likely to over-articulate. Those participants who could not comprehend the sentence and its ambiguities as proficiently could only attempt to imitate the model speaker, but without understanding the importance of the prosodic boundary, may not have emphasized those boundaries.

Miller and Schwanenflugel's (2006) measurement of pausing behavior throughout the whole sentence presented a possible issue given its population. By using young children who are in the midst of developing their reading skills, there is the strong possibility that they will be more naturally disfluent (Lai et al., 2014). They may not have mastered basic reading skills to allow them to reach a level of fluency in which it does not interfere with measuring prosody. The pausing behavior measured by Miller and Schwanenflugel may be more indicative of their participants' overall reading skill level rather than their prosodic fluency, as the pauses analyzed occurred throughout the production of sentences instead of at places which were syntactically or semantically meaningful. This means that the links they drew between good and bad readers and pausing behavior may not be the most accurate representation of the relationship between prosody and reading comprehension.

Additionally, their finding of differences between control readers and poor readers through analyzing F0 changes between declarative statements and yes/no questions may be due in part to their population. Our population was comprised of high school students who have attained a basic grasp of reading skills. The third graders may have shown F0 differences because they were still learning to read, and the ability to understand the F0 differences between those two types of sentences may discriminate between good and bad readers. This may not be the case for our population, as they have all attained a baseline level of reading, they all may recognize the standard difference in final F0 contour between declarative statements and yes/no questions and produce the prosody accordingly. Declarative statements and yes/no questions may more likely be difficult for second grade readers as compared to high school readers, therefore possibly limiting the usefulness of that particular F0 contour for our participants.

The population and stimuli differences found in Miller and Schwanenflugel (2006) also run parallel to those from Binder et al., (2013). Our groups differed significantly in reading comprehension ability while the low literacy skilled adults studied in Binder et al., were deficient in decoding skills. This difference in populations may explain the inconsistencies between our findings; even though both populations are identified as lacking mastery of a particular reading skill, those skills may affect prosody differently. The study by Binder et al., included varied stimuli and measured several types of pitch and pausing behavior, including inappropriate pauses during sentence production. Like the results found in Miller and Schwanengflugel (2006), the readers with low literacy skill made longer pauses and more inappropriate pauses than skilled readers. Although Binder et al., suggested that this relationship was a possible connection between prosody and reading comprehension, but this may not be well-supported. The pauses that were analyzed within the sentences were associated with specific syntactic and semantic boundaries, which strengthened the measurement of a prosodic boundary with comprehension. However, the stimuli were not ambiguous, and therefore unable to more directly test for the presence of prosody to aid comprehension. Furthermore, the deficient decoding skills of the readers may be the main motivator for inappropriate and exaggerated prosody production. Binder et al., did posit that readers who are not skilled decoders may have to allocate more attention to decoding, and therefore have fewer resources to spend on appropriate prosody and comprehension (Walczyk, Marsiglia, Johns & Bryan, 2004). Therefore, the prosody produced could be adversely affected, which may not accurately reflect prosodic skill.

It is also important to note our use of an imitation task for this experiment, and how that affected our results, especially in comparison to previous research. The use of an imitation task contrasts with the methodology of Miller and Schwanengflugel (2006; 2008) and Binder et al., (2012), as they measured the fluency of participants' spontaneous productions of written sentences. Although we measured acoustic features of prosody in this first experiment, because it is an imitation task, it cannot truly measure the prosody produced from spontaneously reading written sentences. Had we had more control over the experimental design and utilized a production paradigm, our results may have been different and more similar to earlier studies.

However, our use of an imitation task may make our results that much more interesting. Participants usually perform well on an imitation task; they can imitate the phonology of a sentence with relative success, especially boundaries (Cole & Shattuck-Hufnagel, 2011). Differences in performance on imitation task may be due to age, with older participants performing better than younger participants, or brain function, as normally developing participants performing better than those with a diagnosed cognitive disorder (Van Der Meulen, Janssen, & Den Os, 1997; Peppé, McCann, Gibbon, O'Hare., & Rutherford, 2007). As our participants were adolescents of comparable ages and were not cognitively impaired, it would be reasonable to expect that there might not be any differences between the control and poor comprehenders in the imitation task. In other words, they should have all performed with similar levels of success. The fact that participants' prosodic cues differed in at least one construct provides possible evidence that reading comprehension skill has a strong enough relationship with prosodic fluency to affect imitation ability. When faced with the ambiguous coordinate structures, poorer reading comprehension skills may have hampered the imitation of key prosodic cues, whereas the control group, with better reading comprehension ability (and a possibly stronger understanding of the role of the prosodic cues in disambiguation) was able to successfully imitate the model sentences.

Experiment 2: Prosodic Sensitivity

Results from Experiment 1 suggest that better comprehenders are more prosodically fluent than poor comprehenders when producing prosodic boundaries in ambiguous coordinate structures. However, these results elucidate only one dimension of the relationship between prosody and reading comprehension. Experiment 2 was designed to test the hypothesis that participants with poor comprehension skills use prosodic cues less effectively than participants with poorer skills. In other words, it is testing whether poor comprehenders are less prosodically sensitive than control comprehenders.

METHOD

Participants

Forty-nine participants contributed to the data analysis of the listening comprehension experiment. Participants were drawn from the same population as the production experiment. Eighteen students participated in both experiment 1 and 2. All were high school students between the ages of 13 and 18. These students were recruited by flyers posted through the community and online advertisement, and they were compensated $25 per hour for their participation. None of these students were diagnosed with learning or reading disabilities. All participants were native speakers of American English and were not exposed to any other languages before the age of 7. Only those students who gave their assent (and had their guardians consent) participated in the study. All procedures were approved by the Yale University Human Investigations Committee. The participants consisted of both specifically poor comprehenders (N = 21) and control participants (N = 28). The two groups of comprehenders were matched on age, gender composition, and phonetic decoding ability (as measured by the Word Attack. All participants scored at or above the normal range of performance for IQ (as measured by Wechsler Abbreviated Scale of Intelligence II). The comprehension groups differed in terms of their reading comprehension standard scores, as measured by the Kaufman Test of Educational Achievement II (KTEA; Kaufman & Kaufman, 2004). Participants classified as specifically poor comprehenders had KTEA standard scores below 95; participants in the control group had KTEA standard scores over 95.

Although two groups of comprehenders were categorized prior, this grouping variable was not used in the statistical analysis; rather reading comprehension (KTEA score) was a continuous variable. The participants differed significantly on all four variables. The parameters for the t-tests are listed in Table 4.

Materials

A total of 84 experimental stimuli were presented to participants. The complete list of stimuli can be found in Appendix B. These stimuli contained pre-recorded sentences from the same female speaker. Twelve practice stimuli were given between the experimental blocks, the responses to the practice sentences were not used for analysis. These practice stimuli followed the same sentence structure of the experimental stimuli in the upcoming block. The pictures used for each construct contained both photographs as well as illustrations, and were all of equal size.

Table 4

Significance Testing between Control and Poor Comprehenders for Experiment 2

	Control	**Control**	**Poor**	**Poor**	**t**	**p**
N	28		21			
Number of Females	13		12			
	Mean	**SD**	**Mean**	**SD**		
Age	15.9	1.6	16.9	1.4	$t(47) = 2.3$	0.026
Word Attack Standard Score	107.25	8.6	102.2	6.9	$t(47) = -2.2$	0.032
IQ (WASI)	111.5	14.9	94.7	8.27	$t(43.7) = -4.9$	<0.001
KTEA Standard Score	112.4	12.6	84.7	5.9	$t(40.3) = -10.2$	<0.001

Accent Disambiguation

The accent location was manipulated in the stimuli for the accent disambiguation construct (AD) to investigate whether participants could use accents to disambiguate between an adjective-noun and a compound noun, such as (12):

(12) Albert grabbed the *top hat* [compound noun] vs. *top hat* [adj.-noun] off the shelf

Farnetani, Torsello, and Cosi (1988) found that participants used accent location, as signaled by pitch, intensity, and duration, to prosodically disambiguate compound words and non-compound phrases. In the adj.-noun condition, the difference between the pitch and intensity of the two words was smaller as both of the words were accented, while only the first word in the compound condition was accented (Farnetani, et al., 1988). The total duration of the words in the compound condition was shorter compared to the overall duration in the adj.-noun condition, indicating a prosodic manifestation of the "cohesion" of the two words in the compound condition.

The manipulation check performed on our prerecorded sentences indicated that there was a significant difference in duration length, mean pitch, and maximum pitch between the compound nouns and the adjective noun phrases. The differences in the prosodic characteristics were analyzed between conditions. Duration was calculated by adding the duration and following silence of the second constituent to the duration and following silence of the first constituent. As the prosodic characteristics of pitch were predicted to differ between the two constituents in each of the conditions, differences were calculated for maximum intensity and mean pitch. The mean pitch of the second constituent in the phrase (e.g. *hat*) was subtracted from the mean pitch of the first constituent in the phrase (e.g. *top*). The same process was followed for maximum intensity. Therefore, the compound nouns were predicted to show a positive difference in the maximum intensity measure, as the first noun is more accented than the second noun. Additionally, the compound noun was expected to exhibit a positive mean pitch difference, as again the first constituent would be accented. Both predictions were supported by the manipulation check. Moreover, the adj.-noun condition exhibited smaller absolute differences between the two constituents than the compound noun condition, as supported by the previous literature. The compound nouns were also found to be significantly shorter in duration length than the adjective noun phrases, displaying more cohesive properties.

The pictures used in this construct were carefully chosen images to correspond to the semantic meaning of the ambiguous words. The compound noun phrase may typically have more standardized visual representations, as it is recognized as a distinct noun within the vernacular, whereas the adjective noun phrase may evoke more varied imagery. It was therefore important to ensure that pictures which represented the adjective noun phrases effectively captured the meaning of the adjective. Such as in Figure 10, where the left-hand image represents the compound noun phrase of a softball (a standardized ball used to play the game of softball) as compared to the adjective noun phrase of a soft ball in in the right-hand image (a ball which appears to be soft). We wanted to be confident that the images chosen in response to the prerecorded sentences were a result of prosodic sensitivity and not due to unclear visual images.

Figure 10. This image depicts the two conditions for the AD construct. The left image depicts the compound noun condition while the right image displays the adjective noun condition for the sentence, *Sally threw a softball/soft ball to her little sister*.

Subject Focus Object Focus

The subject focus object focus (SFOF) construct contained stimuli in which accent location was varied to indicate focus on the subject or the object. For example, in sentence (13),

(13) The rabbit will eat carrots

the accent is either placed on *rabbit* or *carrots* to disambiguate which image in Figure (11) corresponds to the stimulus.

Prior research has shown that accents, in particular contrastive pitch accents, prosodically signify a semantically important word in a sentence; a focus word (Dahan, Tanenhaus, Chambers, 2002; Fraundorf, Watson, Benjamin, 2010; Selkirk, pg. 200, 1986). Contrastive pitch accents refer to a simple pitch rise within a word, as participants have reliably associated this pattern with contrastive information which cues participant attention (Pierrehumbert & Hirschberg, 1990). Focus, as defined by Rooth (1992), "indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions" (cited by Féry & Krifka, 2008, p. 4). In other words, prosodically indicating a focus word signals that there is an alternative to that focus word; the focus word is distinguished because there is another viable option in the context. When the accent is placed on *rabbit*, the image on the left-hand side image is the match. It matches because that image contains two possible subjects (the presence of alternatives), which would necessitate a focus on the subject in the sentence to indicate which option (the squirrel or the rabbit) will eat the carrots. The other image corresponds to an accent (or emphasis) on *carrots*, as two options for the direct object are present in the context (carrots or nuts) which would necessitate the focus on *carrots*.

The stimuli in the subject focus object focus construct were split into the subject focus conditions or the object focus condition for the manipulation check. The two conditions differed in terms of which word in the stimulus was accented. The maximum intensity and the mean intensity of the subjects and objects were recorded in each stimulus in both conditions; accent differences were calculated by comparing those measures of the same word in the two conditions. The differences in accent placement were measured by subtracting the mean intensity of the subject from the mean intensity of the object, with the prediction that the subject focus condition would elicit a negative mean intensity difference and the object focus would result in a positive mean intensity difference. The maximum intensity difference was calculated through the same process, subtracting the maximum intensity of the subject from the maximum intensity of the object. The paired samples t-test revealed significant differences between mean and maximum intensity for the subject focus condition and object focus condition respectively, with the subject focus condition eliciting negative differences (higher intensity on the subject of the sentence) and the object focus condition eliciting positive differences (higher intensity on the object in the sentence).

The images used for the SFOF construct contained visual indication of the subject and the object. In the example above, a red arrow was drawn from the subject to the object, in this case, rabbit to carrots, matching them to the prerecorded sentence. Each image contained either a realistic alternative for the subject or the object of the sentence. The alternative was related to the subject or the object and semantically fit with the prerecorded sentence. The alternative for the subject rabbit from the example was another small mammal (a squirrel) which could also be reasonably assumed to eat carrots.

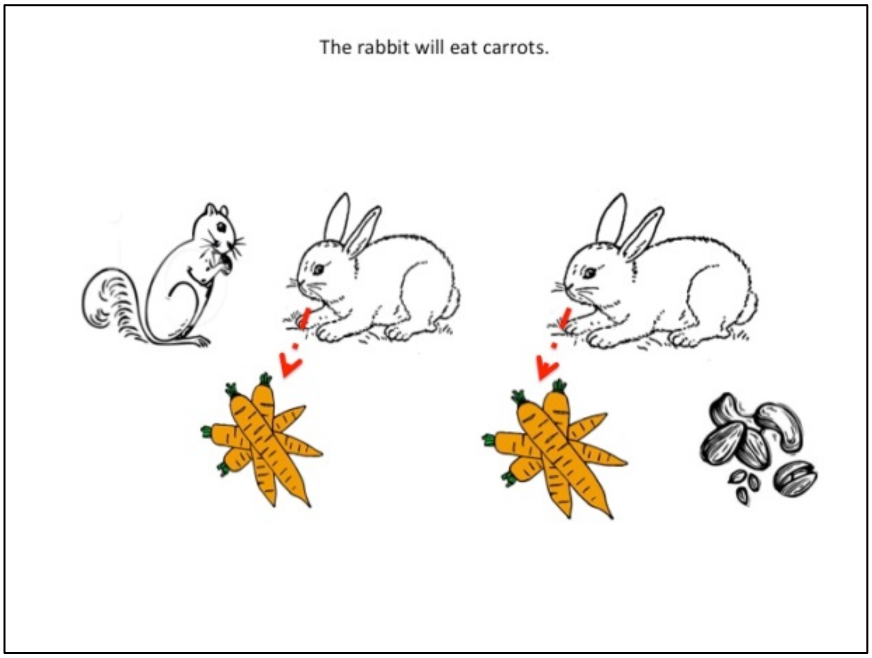

Figure 11. An example of an image displaying the two possible choices for the SFOF construct for example sentence (13). The left image depicts the subject focus, the right image depicts the object focus.

Contour Ambiguity

Contour ambiguity (CA) stimuli manipulate both accent location and intonational contour in order to disambiguate the correct interpretation. Each stimulus began with the phrase "it looks like…" and finished with a singular noun, such as (14):

(14) It looks like a backpack

based on stimuli from Kuramada, Brown, and Tanenhaus (2012). For the example earlier of "it looks like a backpack", Figure 12 contains the two images shown.

The intonational contour of the sentence is interpreted as either the affirmative "it looks like a backpack… (and it is a backpack)" to correspond with the image on the right, or "it looks like a backpack… (and it is not a backpack)" which corresponds with the image to the left. The affirmative interpretation occurs when the verb "looks" is read with a slight pitch rise, while the negative interpretation has an elongated "looks" with a fall-rise (a pitch change associated with the ending of a question). Additionally, the direct object of the stimulus was found to be produced with a pitch fall in the affirmative condition, and a pitch rise in the negative. Kurumada (2012) found that participants disambiguated the correct interpretation using the prosodic contours in the stimuli, such that they were more likely to correctly choose the corresponding target picture to the prerecorded sentence when the prosodic cues were evaluated as unambiguous. Sentences with ambiguous prosodic cues, meaning they were not salient, significantly impaired accuracy of performance.

The stimuli of CA were differentiated in the manipulation check through the final pitch contour, with the affirmative condition predicting a pitch fall for the direct object (DO) of the stimulus while the negative condition elicits a pitch rise. Pitch slope for the DO was calculated by subtracting the average pitch recorded at 20% completion of the direct object production from the mean pitch at 80% of completion. This would result in the affirmative condition producing a negative mean pitch slope while the negative condition would produce a positive mean pitch slope. The paired samples t-test revealed a significant mean difference between the negative condition and the affirmative condition, with both conditions producing the predicted pitch slopes. The pitch slope for the verb "looks" was calculated by subtracting the mean pitch recorded at 20% completion of production of the verb from the mean pitch recorded at 40% of completion. As the verb was only one syllable, there were fewer pitch intervals at which to record, which is why 40% of word completion is used as the "ending" percentage of completion. The pitch slope for the affirmative condition was predicted to be significantly smaller than the pitch slope for the negative, as the negative condition contains a strong pitch fall-rise whereas the affirmative verb pitch rise is very small. The production of the verb supported these predictions and attained a significant difference.

The images chosen for the CA construct had to contain two images which could correspond to the direct object of the sentence. One image would be used to represent the affirmative condition, in which the object in the image would be familiar and contain well-known characteristics of the direct object. Such as in

Figure 12, the object on the right corresponds with the affirmative condition as it is more easily recognizable as a backpack than the image on the left. The left-hand side image references the negative condition because it contains some distinguishable elements of the direct object (in this case it has two shoulder straps like a conventional backpack), but still retains more unfamiliar characteristics (such as its shape). The images chosen have some resemblance of one another, while still maintaining clear cut differences between conventional and the unfamiliar.

Figure 12. An image containing two possible choices for the CA construct pertaining to the example sentence (14). On the left depicts the choice for the negative interpretation, while the one on the right depicts the affirmative.

Coordination

The stimuli used in the coordination (CO) constructs contained syntactic attachment (height of attachment) ambiguities. This construct is motivated by the same research as the two-one and two-two disambiguation by phrasing construct from the imitation task experiment. The sentence (15) can be disambiguated by two-one phrasing:

(15) Sarah and Will or Molly paid for dinner

(15a) Sarah and Will || or Molly paid for dinner

Or through two-two phrasing:

(15b) Sarah || and Will or Molly paid for dinner

The two-one (15a) communicates that Sarah and Will paid or only Molly paid, while the two-two phrasing (15b) indicates that Sarah and Will paid for dinner or Sarah and Molly paid for dinner. These interpretations are prosodically disambiguated through phrase boundaries. As presented earlier, when a series of subjects are listed, especially with the use of the word “or”, groupings will occur (Wagner, 2005). The groupings that are formed are prosodically signaled through boundaries, and end before a conjunct that is not attached to the preceding phrase. These boundaries can be marked by the lengthening of the phrase-final word and a subsequent pause. It is important to note that the constituents that were ambiguously grouped did not exceed lengths of two syllables. Prior research has shown that participants take into account constituent length when placing prosodic phrase boundaries, and so controlling for the length of the constituents ensures the placement is due to syntactic disambiguation (Clifton, Carlson, & Frazier, 2006).

The manipulation check of the prerecorded sentences for the CO construct utilized the prosodic characteristic of duration as the difference between constituents in each condition. In the two-two disambiguation, we predicted that a prosodic boundary will be placed after the first conjunct (*Sarah*) in the sentence, while in the two-one disambiguation a boundary will be placed after the second conjunct (*Will*). For our analysis, the duration plus following silence of conjuncts 1 and 2 were compared across the two disambiguation conditions. Specifically, we expected the duration and silence of first conjunct to be significantly longer in the two-two disambiguation as compared to the two-one disambiguation. We held the opposite prediction for second conjunct : the duration and silence would be significantly longer in the two-one disambiguation than the two-two.

A paired samples t-test was performed for both conjuncts. The first conjunct showed a significant difference in duration between the two-two and the two-one disambiguation, with the two-two condition eliciting a longer duration and silence than the two-one condition. The duration and silence for second conjunct differed significantly between conditions, with the two-one disambiguation eliciting a longer boundary than in the two-two condition. For this construct, the images used for each condition visually grouped the constituents.

Figure 13, which corresponds to the previous example (15), displays the image on the left-hand side which matches the two-two disambiguation (15a), while the image on the right matches the two-one condition (15b). The grouping of the visual constituents was salient and visually connected to an image depicting the main verb, like in the example (15) where the constituents are grouped by who is paying for dinner which is represented using an image of cash money and a restaurant bill.

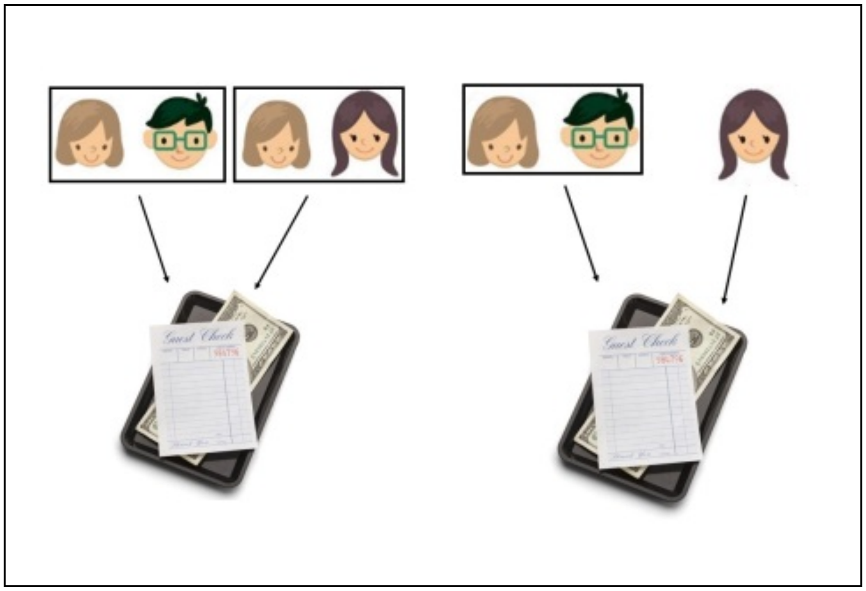

Figure 13. An example depicting two possible choices for the CO example sentence (15). The image on the left depicts the two-two condition, while the image on the right depicts the two-one condition.

Phrasing Ambiguity

Phrasing ambiguity (AP), like CO, also used syntactic attachment as the catalyst for prosodic manipulation, however within this construct the ambiguity was between an instrument and modifier interpretation. Motivated by Snedeker and Trueswell (2003), this construct utilized an ambiguous sentence such as (16), which is realized either in an instrument condition (16a) or a modifier condition (16b):

(16) Tap the frog with the flower

(16a) Tap the frog || with the flower

(16b) Tap || the frog with the flower

The instrument condition instructs the participant to use the flower to tap the frog, while in the modifier condition it signals that in that there is only one frog with a flower and that is the frog to tap. Figure 14 visualizes both interpretations, with the modifier interpretation (16b) on the left, and the instrument on the right hand side (16a). Snedeker and Trueswell (2003) found that when faced with an ambiguous presentation, such as having access to a flower to use to tap an object and seeing a frog with a flower, listeners will use prosodic cues to disambiguate the instructions. Those cues manifested in pauses and duration, which resulted in a boundary before “with” and a lengthening of the direct object (“frog”) in the instrument condition (16a) and a lengthening of the verb “tap” in the modifier interpretation (16b). Therefore the instrument condition (in conjunction with the boundary), elicited a longer prepositional phrase than the modifier condition.

Since the AP construct utilized duration as its signifier for the prosodic differences within each condition that was the prosodic characteristic used for the manipulation check. This construct contained durational differences before and at the beginning of the prepositional phrase. The instrument condition was predicted to have a lengthening of the direct object, coupled with a prosodic boundary between the direct object and the prepositional phrase, resulting in an overall longer verb and prepositional phrase duration as compared to the modifier condition. The measurement of duration was calculated by adding the duration and following silence of the direct object to the total duration and silence of the prepositional phrase. The paired samples t-test revealed a significant difference between the duration of the instrument condition and the modifier condition, with the instrument condition eliciting a longer duration than the modifier condition.

The AP construct consisted of images which represented both the modifier and instrument conditions. As shown in Figure (7), the same smaller individual images (i.e. the hand, the flower and the frog) were used to compose both condition images, but were positioned accordingly to reflect the semantic differences. In the instrument condition, the instrument was placed between the hand and the direct object, while in the modifier condition the same image used as the instrument was placed on the direct object ensuring it did not overlap with the hand.

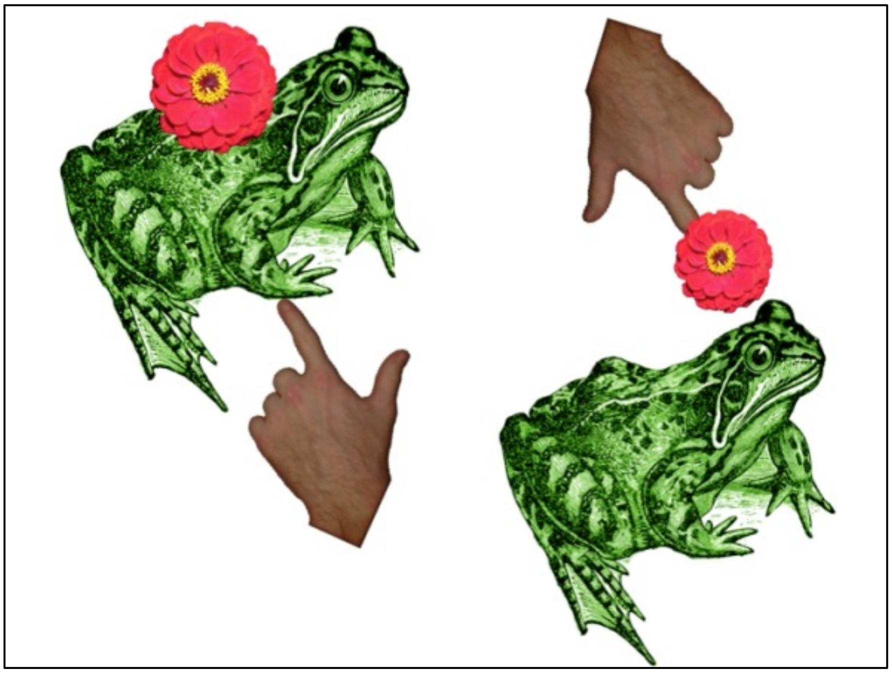

Figure 14. Two possible choices for the AP construct for the previous example (16). The left image displays the modifier condition, while the right image depicts the instrument condition.

Disjunctive Questions

Disjunctive questions (DQ) are questions which include a yes/no interpretation or an alternative interpretation (i.e. asking to choose one option). The example question (17) asks if Paula only sings or if she only dances (alternative question), or it asks if she either sings or dances *or* does neither (yes/no question) (Pruitt & Roelofsen, 2013).

(17) Does Paula sing or dance

Participants in the Pruitt and Roelofsen (2013) study relied heavily on the final pitch contour (beginning with "or" and including the final constituent in the question) to disambiguate between the two questions: an alternative question does not include a final pitch fall, while a pitch-rise at the end of the question was more likely to elicit a yes/no interpretation from participants.

Disjunctive questions are disambiguated through the pitch slope of the final constituent, and so we analyzed those pitch slopes for the manipulation check. We calculated the mean pitch slope by subtracting the mean pitch recorded at 40% of the constituent completion from the mean pitch recorded at 70% of completion. These calculations would result in a negative mean pitch slope for the alternative condition and a positive mean pitch slope for the yes/no condition. The paired samples t-test revealed a significant difference between mean pitch slopes of the alternative and yes/no conditions, with the alternative condition producing a negative mean pitch slope (final pitch fall) and the yes/no condition producing a positive mean pith slope (final pitch rise).

The images used in the disjunctive question construct utilized the no symbol pictogram, a universally recognized image indicating “no”, to symbolize the possible “no” response which could be elicited from a yes/no question. As seen in Figure 15, this image depicts the “yes” option on the left hand side of the slash mark in the yes/no question condition of *Does Paula sing or dance*. All images indicating the options in the yes/no question condition were formatted in the same order. The picture indicating the disjunctive question condition, in which Paul either sings or dances is presented in Figure 16. The placement of the options in the disjunctive question followed the same order: the first option (e.g. *sing*) is presented on the left of the slash mark while the second option follows the slash mark. The pictures utilized both illustrations as well as photographs.

Figure 15. Image depicting the yes/no choice for the DQ construct for the previous example question (17). The left images include both possible options within the question (e.g. *sing* or *dance*) which depicts the *yes* portion of the response, while the right image displays a forbidden symbol which indicates the *no* portion of the yes/no response. The slash in between separates the images into their respective categories of *yes* and *no* for the yes/no condition.

Figure 16. An image depicting the alternative question choice for the DQ construct for the previous example question (17). The image on the left displays one of the options from the question (*sing*) while the image on the right displays the other option (*dance*). The slash mark between the two indicates that these images depict two separate options in the alternative condition (e.g. she either sings *or* dances).

Manipulation Check

As this experiment utilized a listening comprehension task, in which results were contingent on the prosodic differences perceived by participants', it was necessary to ensure the presence of those prosodic characteristics. The audio stimuli were recorded by a native English speaker through a Shure SM10A head-mounted microphone, connected to a Rolls Mini-Mic pre-amplifier. Productions were recorded with a sampling rate of 44100 Hz. The waveforms were force-aligned with the sentences using the ProsodyLab Aligner (Gorman, et al, 2011) and acoustic features were extracted using scripts implemented in Praat (Boersma & Weenink, 2011). Each stimulus in every construct contained two possible interpretations or conditions, therefore paired samples t-test were performed on the audio stimuli to explore the prosodic differences each condition. The analyses specifically tested the acoustic features which were predicted to differ based on interpretation; those predictions, in turn, were originally based on prior psycholinguistics work. The results of these analyses (means, *t*-values and *p*-values) are presented in Table 5.

Table 5

Parameters from Paired Sample t-tests for Manipulation Check

Construct	Condition 1 (Mean)	Condition 2 (Mean)	SE	t	p
AD	**Adj/Noun**	**Compound**			
Max Intensity	3.41	4.49	2.32	3.43	0.004
Mean Pitch	-37.88	62.84	15.47	6.51	< 0.001
Duration	0.94	0.7	0.04	-6.91	<0.001
AP	**Instrument**	**Modifier**			
Duration	1.48	0.96	0.08	-6.94	<0.001
CA	**Affirmative**	**Question**			
Pitch Slope "Looks"	-2.84	26.95	4.76	-6.26	<0.001
Pitch Slope Direct Object	-138.87	70.44	27.87	-7.51	< 0.001
CO	**Two-One**	**Two-Two**			
Duration First Conjunct	0.44	0.67	0.05	-2.43	0.029
Duration Second Conjunct	0.73	0.47	0.05	2.27	0.040
DQ	**Alternative**	**Yes/No**			
Pitch Slope	-64.30	82.71	48.04	3.06	0.010
SFOF	**Object Focus**	**Subject Focus**			
Mean Intensity	1.42	-10.19	0.73	-15.92	<0.001
Max Intensity	5.02	-9.55	0.73	-20.1	< 0.001

Testing Materials

The materials used for testing prior to the prosodic sensitivity experiment are outlined in detail in the method section for the imitation task in experiment 1. Participants provided sensory and demographic information either before the imitation task or before this current task. They also completed the Word Attack test from the Woodcock Johnson, the reading comprehension section of the KTEA, and the Perceptual Reasoning Index (PIQ) of the WASI-II if they did not participate in the imitation task.

Procedure

Participants, unless having previously participated in the production experiment, completed the demographic questionnaire as well as three measures of standardized testing: Word Attack subset of the Woodcock Johnson, the WASI, and the KTEA. Once gaining the informed consent of the participant, the participant was seated in front of a BenQ monitor with screen resolution of 1024 x 768. The screen, on which the visual stimuli would appear, was positioned approximately 95cm from the participant's eyes. Eye movements for this experiment were monitored using an Eyelink 1000 Plus eye-tracker, with sampling at 1000 Hz. A chin rest was used to reduce movement during the experiment.

Before experimentation, participants were instructed that two images would appear side-by-side on the screen for each pre-recorded sentence, and they must choose the one image which best matched the stimulus.

The participants were permitted to repeat any of the sentences by clicking on a fixation cross which was displayed in between the two images on the screen. The stimuli were run in a total of 6 blocks, with approximately 14 trials in each. Each block of stimuli corresponded with a specific prosodic construct (AD, AP, SFOF, CA, CO, DQ). The presentation of each block was randomized and the trials within each block were also randomized. Between the experimental blocks, participants were presented with two practice trials which included brief reiterations of the directions. These practice trials were used to ensure participant comprehension and allowed the experimenter to clarify any misunderstandings. The experimenter would illustrate the contrastive stress present in the sentence and how that stress relates to each image.

An example of this feedback dialogue is represented in (18):

(18)

(Pre-recorded sentence): "Jerald used redwood to make a bookcase."

Experimenter: "With this sentence, because 'redwood' is said as if it is one word, you would choose the picture showing a type of wood instead of a piece of wood that is literally red."

(Pre-recorded sentence): "Jerald used red wood to make a bookcase."

Experimenter: "With this sentence, because red and wood are separated and "wood" is stressed, you would choose the picture of a piece of wood that is literally red."

The experimenter would also give accuracy feedback on the participant's choice, after which the participant had the opportunity to amend the original image selection. Participants were not given feedback by the experimenter during the experimental blocks. The experimenter gave the same amount of feedback (in terms of details and length) for all participants for all of the practice examples across the 6 blocks.

RESULTS

The final inclusion of the 6 constructs resulted in a total of 86 experimental stimuli, with 1 trial per stimulus for all 38 participants which created 3268 trials total. The prosodic sensitivity task was a within-subjects design. Each participant heard every sentence once in only one of its two conditions. The order for each stimulus and its condition was randomized, with approximately equal division between the two disambiguations in each construct (i.e. 6 stimuli would be presented in the two-one disambiguation condition for CO while 7 stimuli would be in the two-one disambiguation condition). The data collected from this prosodic sensitivity experiment were accuracy data, which was categorized as either a correct or incorrect choice. Due to the categorical nature of the data, a binomial general linear mixed effects model of regression (logistic regression model) was utilized to predict the accuracy data.

The goal of this experiment was to better understand how prosodic sensitivity affects listening comprehension performance, and how reading comprehension ability accounted for variability in that performance. The sentence types used in this experiment are diverse, with each type containing a particular set of prosodic characteristics that have been shown in psycholinguistic research to disambiguate ambiguous stimuli. Our statistical analyses were conducted to test whether participants were sensitive to the prosodic cues and if they successfully utilized them to aid in choosing their answer, and whether the poor comprehenders performed significantly worse than the control group. Overall, there is a strong positive correlation between comprehension ability and accuracy performance (*r* = 0.47), which is plotted in Figure (18).

This analysis was calculated on a trial-by-trial basis across participants for each individual construct, meaning that each model analyzed approximately 551 accuracy scores. We treated each construct and its accuracy scores as independent models. Each model was run with the behavioral factors of comprehension score KTEA, IQ, age, as well as random intercepts for both participant and item. Every model included the measures of IQ and age, as both of those factors may be generally relevant to our measures, and we wanted to ensure that the predictive ability of the KTEA scores were not including variance accounted for by IQ or age. Of the 6 constructs modeled, only 2, disjunctive questions and coordination ambiguity, contained significant prediction models based on KTEA scores, with an alpha level of 0.05. The full parameters of the models can be found in Table 6. Overall mean and std. deviation values of the accuracy performance of each construct across participants are displayed in Table 7.

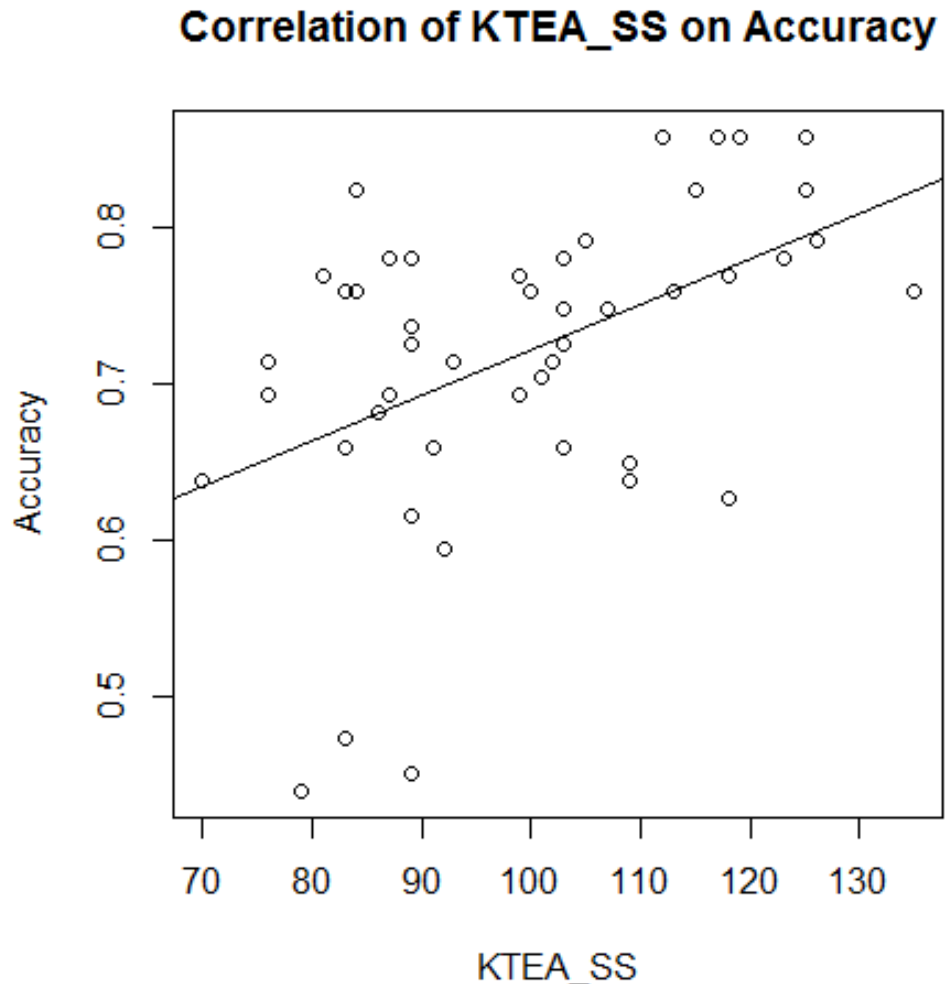

*Figure 18***.** Scatter plot of the correlation of reading comprehension score (KTEA_SS) on Accuracy for experiment 2. The regression line estimated: y' = 0.429 + 0.002(KTEA_SS), with both the intercept (SE = 0.082, t = 5.215, $p <$ 0.001) and slope attaining significance (SE = 0.001, t = 3.567, $p <$ 0.001).

Table 6

Parameter Estimates of the General Linear Models for Experiment 2

Construct		Est.	SE	z	p
AD					
	Intercept	-4.65	2.58	-1.80	0.072
	KTEA	0.03	0.02	2.12	0.034 *
	IQ	0.02	0.02	1.02	0.307
	Age	0.06	0.12	0.54	0.589
AP					
	Intercept	-4.06	1.96	-2.08	0.038 *
	KTEA	0.02	0.01	1.55	0.120
	IQ	0.01	0.01	0.45	0.650
	Age	0.10	0.09	1.09	0.278
DQ					
	Intercept	-9.94	4.22	-2.36	0.019 *
	KTEA	0.04	0.03	1.42	0.156
	IQ	0.03	0.03	1.12	0.262
	Age	0.35	0.19	1.87	0.062
CA					
	Intercept	-3.77	4.71	-0.80	0.424
	KTEA	0.02	0.03	0.79	0.432
	IQ	0.02	0.03	0.80	0.422
	Age	0.18	0.22	0.82	0.415
CO					
	Intercept	-3.23	2.39	-1.35	0.176
	KTEA	0.04	0.01	2.37	0.018 *
	IQ	0.00	0.02	0.01	0.989
	Age	0.01	0.11	0.10	0.923
SFOF					
	Intercept	-1.42	2.93	-0.49	0.627
	KTEA	0.00	0.02	-0.27	0.784
	IQ	0.02	0.02	1.00	0.318
	Age	0.07	0.14	0.48	0.628

Note: * indicates estimated significance beyond the .05 level

Table 7

Means and Standard Deviations for Overall Accuracy by Construct

Construct	Mean Accuracy	Std. Deviation
AD	0.76	0.42
AP	0.53	0.50
CA	0.87	0.34
CO	0.62	0.49
DQ	0.82	0.38
SFOF	0.70	0.46

Accent Disambiguation (AD)

The stimuli in the AD construct contrasted two interpretations of a noun phrase; such as in the example noun phrase *top hat*, which could be construed in context as either meaning a formal dress hat, or a hat which is placed on the highest shelf. The prosodic features which have been shown to disambiguate the interpretations include the pitch, intensity and duration of the words of interest. The general linear model predicting accuracy scores based on IQ, age, KTEA score, and random intercepts for both participant and item revealed a significant effect for KTEA as a predictor ($z = 2.12$, $p = 0.034$). This indicates that the participants' reading comprehension ability may play a role in predicting their performance on prosodically disambiguating stimuli based on accents.

Subject Focus Object Focus (SFOF)

This construct contained stimuli which were interpreted with either a focus on the subject or the object for successful disambiguation based on the visual stimuli presented (e.g. *The rabbit ate the carrots*). This focus would manifest by aiding the listener by indicating that the focus word is emphasized because the context provides an alternative; each image presented either two options for the subject (a rabbit and a squirrel) or the object (a carrot or nuts). Prosodically, the differences in the conditions were heard through pitch accent placement, such as intensity. The general linear model predicting accuracy did not find a significant effect for the use of KTEA scores as a predictor ($z = -0.27$, $p = 0.784$), after accounting for age, IQ, and random intercepts for participant and item.

Contour Ambiguity (CA)

These stimuli in the CA construct consisted of sentences which began with *it looks like…* and concluded with either an affirmative interpretation or a negative questioning interpretation. These two intpretations differed in both their accent location as well as their prosodic contour. The general linear model which analyzed the predictors of KTEA score, IQ, age, and random intercepts for participant and item revealed a non-significant effect of KTEA score on accuracy prediction ($z = 0.79$, $p = 0.432$). Comprehension ability does not appear to mitigate prosodic sensitivity in accuracy scores for this particular construct.

Coordination (CO)

These stimuli within CO were syntactically complex sentences which included two interpretations for height of attachment (e.g. *Sarah and Will or Molly paid for dinner*). Prosodic phrase boundaries were utilized in the disambiguation of these stimuli, which presented themselves in terms of durational differences in phrase-final words (*Sarah* or *Will* depending on the condition). The general linear model which predicted accuracy scores using KTEA scores, age, IQ and random intercepts for participant and item found that KTEA significantly predicts accuracy scores ($z = 2.37$, $p = 0.018$). This result presents evidence that reading comprehension skill may predict prosodic sensitivity ability, specifically for the prosodic characteristic of duration, affecting accuracy scores for the CO stimuli.

Phrasing Ambiguity (AP)

The construct of AP contained stimuli which differed on a modifier or instrument interpretation of its stimuli (e.g. *Tap the frog with the flower*). The two conditions were prosodically characterized by differences in pausing and duration, specifically preceding the prepositional phrase. The general linear model did not present an effect of KTEA scores on the prediction of accuracy for AP ($z = 1.55$, $p = 0.120$), after accounting for age, IQ, with the inclusion of random intercepts for participant and item. This indicates that reading comprehension ability may not modulate accuracy performance on this construct task.

Disjunctive Questions (DQ)

Disjunctive questions (e.g. Does Paula sing or dance) are questions which include either a yes/no interpretation or an alternative interpretation. The two conditions are disambiguated through the pitch-final contour of the question, with the final constituent in yes/no questions inducing a pitch rise while a pitch fall occurred in alternative interpretations. The general linear model used to predict accuracy scores using KTEA scores, IQ, age, and random intercepts for participant and item found that KTEA does not significantly predict accuracy score ($z = 0.79$, $p = 0.432$). Reading comprehension ability appears to not have a relationship moderating the accuracy performance of this construct.

Correlation Analysis Between Experiments 1 and 2

Two constructs between the two experiments utilized the same prosodic manipulation across identical conditions. The ambiguous coordinate in experiment 1 and the coordination construct in experiment 2 contained stimuli with prosodic phrase boundaries for disambiguation of grouping. The pre-recorded sentence *Ann and Bobby or Nancy will come* occurred with a prosodic phrase boundary after either the first conjunct (*Ann*) for the two-two disambiguation or after the third (*Bobby*) which resulted in the two-one disambiguation. This boundary prosodically manifested in a lengthening of the duration of the preceding constituent and longer silence following its completion. In the imitation task experiment, the mixed-linear effects general regression model analyzed the condition (two-two vs. two-one), the comprehension group (control vs. poor readers), and the interaction of the two factors on the duration of each conjunct (the first or second). Both constituent models showed significant or near significant effects of each variable for predicting duration, however only the second conjunct analysis elicited a significant interaction. As noted earlier, the equivalent construct in the prosodic sensitivity experiment (CO) found that reading comprehension score significantly predicted accuracy performance above and beyond IQ and age within a general linear model.

Both of the constructs within their respective experiments produced promising results, so a correlation analysis was run to analyze the relationship between the durations of first conjunct and 3 in each condition in the imitation task, the accuracy rates in the prosodic sensitivity task, and reading comprehension scores. A Pearson's correlation analysis was run on a participant-by-participant basis. Only participants who had participated in both experiments were included, leaving our participant pool at 18. Accuracy rates were calculated for each participant by averaging their accuracy scores from each of the 15 trials in the CO construct in the prosodic sensitivity experiment, creating a value ranging from 0 to 1, with a higher value indicating better performance. The values used in the analysis from the imitation task were mean durational difference for both constituents for every participant. This was calculated by first averaging the durations of the first and second conjuncts (separately) in both conditions across the three productions per stimulus. Then the durational difference between conditions was calculated for each specific conjunct (1 and 2) for each participant; the difference was found between the duration of first conjunct (*Ann*) in *Ann || and Bobby or Nancy will come* (2-2 condition) and the 2-1 condition (*Ann and Bobby || or Nancy will come*). More specifically, the duration of the conjuncts of 2-2 were subtracted from their respective sentence duration of the conjuncts of the 2-1 condition before averaging across stimuli. This specificity allowed for better control over both individual differences as well as differences that were possibly present between stimuli. Once the durational differences were found for every conjunct (first and second) within their own individual sentences, those differences were averaged across for each participant. Each participant then ended up with an average durational difference between conditions for conjuncts 1 and 2, which we used in the correlation analysis. This lead to specific expectations for trends, in that we expected the average durational difference for second conjunct to be a positive value, as second conjunct produced longer durations in the two-one condition. The opposite expectation was held for the first conjunct.

The durational difference for second conjunct was correlated with the accuracy rating, which yielded a Pearsons' r of 0.102 and a non-significant *p* value of 0.687. The durational difference for first conjunct was also not significant with an r value of 0.127 and a *p* value of 0.616. Neither correlation indicates a strong relationship between the prosodic productions of the conjuncts in experiment 1 and the accuracy performance in experiment 2.

DISCUSSION

From the six total constructs that were analyzed for the prosodic sensitivity experiment, the accuracy performance of two constructs, accent disambiguation and ambiguous coordinate structures, were significantly predicted by reading comprehension ability. For the logistic regression analyses conducted in this experiment, this means that reading comprehension scores significantly predicted accuracy performance above and beyond age and IQ of the participants. The ambiguous coordinate structure construct had a significant relationship with reading comprehension ability in both experiments. This is not surprising, as prior research has shown that duration and pausing behavior are well documented characteristics of prosodic phrase boundaries for signaling height of attachment (Kjelgaard & Speer, 1999; Snedeker & Trueswell, 2003). In order for a participant to correctly choose the corresponding image for the disambiguation condition, the participant would have to first detect the prosodic phrase boundary in the stimulus and then successfully comprehend its role in disambiguation. The participant must first understand the role of the prosodic cue (i.e. how it changes the semantic structure) in order to correctly utilize it for comprehension purposes. In this way, it is possible that participants who are skilled comprehenders are more sensitive to prosodic cues because they understand their importance in supporting comprehension.

The significant relationship between reading comprehension ability and the accent disambiguation construct is an interesting finding, both for our study and for prior work. The two conditions within this construct, adjective-noun phrase or compound noun, differed in terms of F0, intensity and duration (Farnetani, et al., 1988). Therefore, detecting the semantic relationship between the target words may have been complex, as it possibly involved identifying and understanding several prosodic cues. Although there was no construct to sufficiently parallel the accent disambiguation construct in the imitation task, the use of duration and following silence as indication of a prosodic break is, however, familiar. The results from this construct strengthen the relationship between detecting prosodic boundaries and reading comprehension ability.

Although prior psycholinguistic research has studied the accent differences between adjective-nouns and compound nouns, it has not been studied in relation to reading comprehension skill and prosodic sensitivity. The role that reading comprehension may play in affecting the accuracy rate for this construct may be due to the sophistication of disambiguating the stimuli. In other words, the accent disambiguation construct may have proved challenging to our participants. There are several prosodic cues associated with disambiguating adjective noun phrases and compound nouns, and that disambiguation is not always reliably produced by the same prosodic cues in the same pattern (Plag, Kunter, Lappe, & Braun, 2008). This difficulty may have been a factor in why reading comprehension differences mediated accuracy performance for our participants. According to results found by Benjamin and Schwanenflugel (2010), participants rely on prosody (specifically producing appropriate prosody) to facilitate comprehension in harder sentences. This same reliance on prosody may be reflected in prosodic sensitivity. Those participants who are skilled comprehenders may be defter in recognizing prosodic cues and can therefore utilize them for comprehension and accurate image selection. The control group may also be better at recognizing a wider range of prosodic cues than the poor comprehenders, which would ultimately facilitate the disambiguation of these complex stimuli.

Although psycholinguistic research has provided support for the presence of prosodic cues found in ambiguous coordinate structures and accent disambiguation stimuli, as well as their use in disambiguation, studies exploring the relationship between prosodic sensitivity and reading comprehension have not tested specific elements of our stimuli. Neither prosodic phrase boundaries nor the characteristic of duration have been examined in previous research. Pitch features are virtually nonexistent in the prosodic sensitivity litearture. The focus of the prior research has consisted of almost exclusively studying stress patterns within sentences (Clin, Wade-Woolley, and Heggie, 2009; Holliman, Sheehy, and Wood, 2008; Whalley & Hansen, 2006). These studies have found support that sensitivity to the prosodic characteristic of stress correlates with reading comprehension ability. Some studies have tried to isolate specific components of prosodic sensitivity for study, such as Holliman, et al., (2014). Their study included the features of intonation, stress and timing at various levels (word, phrase, and sentence) and found that the different prosodic features moderately accounted for variance in the relationship between prosody and reading comprehension ability. Our study, with both more precise prosodic measurement and well documented semantic manipulations for our stimuli, was better able to detect those differences between prosodic characteristics and reading comprehension ability.

GENERAL DISCUSSION

The purpose of the current study was to propose a unified and quantifiable definition of prosodic fluency, and to explore not only the relationships between prosodic fluency and reading comprehension and prosodic sensitivity and reading comprehension, but how those relationships may fit with one another. Prosodic fluency, as earlier defined, is the ability to produce appropriate prosody in terms of the prosodic characteristics produced and the placement of those characteristics (which is dependent on semantics and syntax). Prosodic sensitivity is the ability to successfully detect and subsequently utilize prosodic cues from audio stimuli.

Of the several syntactic and semantic structures tested throughout the prosodic fluency and the prosodic sensitivity experiments, only one construct and prosodic characteristic was paralleled in the two experiments. In the prosodic fluency experiment, the difference in the duration of the second conjunct produced in the ambiguous coordinate structure differed significantly across both condition (two-one and two-two) and reading comprehension groups. The significant interaction showed that participants with better reading comprehension skills produced a larger difference in the duration produced between conditions, while those with poor reading skills produced a smaller difference. Furthermore, the accuracy performance on the same structure (CO) in the prosodic sensitivity experiment was found to be significantly predicted by reading comprehension ability. Participants with higher reading comprehension ability performed better on the listening comprehension task as compared to those with lower ability.

Although the correlation analysis between this construct's results in the two experiments proved to be non-significant, the significance of each of these individual experiments may elucidate a connection between the skills of prosodic sensitivity, prosodic fluency and reading comprehension.

Both of these constructs utilized the same manipulation, constituent grouping conditions, and relied on the same prosodic characteristic, duration and subsequent pausing, to disambiguate the syntactic structure. These aspects make up the important thread when discussing the possible relationship between prosodic fluency, prosodic sensitivity and reading comprehension. These aspects also prompt certain questions. Why did this particular syntactic structure elicit significantly strong effects of prosody? Why is duration the only prosodic characteristic associated with the significant constructs between experiments? Part of the answer may lie in the strength of these components; both prosodic phrase boundaries due to grouping and the use and production of durational cues are well supported in the literature.

The stimuli of the ambiguous coordinate structure were created to elicit particular prosodic phrase boundaries to disambiguate the height of syntactic attachment ambiguity. These stimuli were ambiguous, in that the sentence *Ann and Bobby or Nancy came* could have the following syntactic structures based on semantic grouping:

(19) Ann || and Bobby or Nancy came

(20) Ann and Bobby || or Nancy came

The ambiguity lies in the attachment of the constituent after *or,* and whether or not it attaches with the first conjunct. The use of the conjunction *or* denotes alternatives; in these stimuli it stipulates whether the alternatives are *Ann* and *Bobby*/*Ann* and *Nancy* or *Ann* and *Bobby*/*Nancy.* The prosodic break after the first conjunct (*Ann*) denotes an early closure (height of attachment), while the prosodic break after the second conjunct is a late closure. The closure refers to where in the sentence (early or late) the prosodic phrase is ended, which is thus marked by a prosodic cue. Prior psycholinguistic research has shown strong support for both the disambiguating strength of prosodic phrase boundaries, as well the strength of prosodically grouping meaningful constituents. For our study, prosodic phrase boundaries in the constructs were measured through duration and following silence of the word preceding a boundary; however, prosodic phrase boundaries can be measured by duration, pausing, and pitch features.

Prosodic phrase boundaries are a well cited prosodic feature in terms of its effectiveness in disambiguation. These boundaries can be placed based on semantic relationships between words, which can be helpful to break the sentence into meaningful groups for easier comprehension (Breen et al., 2011). Additionally, other studies have shown that prosodic phrase boundaries facilitate comprehension through disambiguation of syntactic attachment (Kjelgaard & Speer, 1999; Snedeker & Trueswell, 2003). Prosodic phrase boundaries have been both produced by speakers and utilized by listeners in the prior studies, demonstrating probable salience in prosodic fluency and prosodic sensitivity of participants.

The specific syntactic structure within our stimuli, the grouping of constituents with the conjunction *or*, has been shown to elicit prosodic phrase boundaries. In a prosodic sensitivity study (Streeter, 1978), listeners had to determine the grouping of ambiguous algebraic expressions (e.g. [*A plus E] times O* or *A plus* [*E times O]*) based on prosodically cued boundaries. Both the speakers and participants reliably utilized the prosodic characteristic of duration, as the speakers lengthening of the constituents preceding the grouping boundary significantly improved listeners' accuracy of grouping judgement. Furthermore, adult listeners' have been shown to be cognizant of the reasons behind prosodic breaks of sentences with ambiguous grouping of constituents with conjunctions, with an understanding that it can convey disambiguating syntactic information (Clifton, Carlson, & Frazier, 2006).

The prosodic characteristic of duration, as well as pausing behavior, is widely measured as a prosodic characteristic in both psycholinguistic studies and those studying prosody and reading comprehension. Several studies exploring prosody and reading comprehension utilized pausing behavior to indicate that readers with poor comprehension ability make more inappropriate pauses than those who are skilled (Binder et al., 2013; Miller & Schwanenflugel, 2006; Schwanenflugel, 2004). Psycholinguistic research has also measured duration as a boundary cue, with the important note that both speakers and listeners utilize durational cues for comprehension (Breen et al., 2011; Clifton, Carlson, & Frazier, 2006; Kjelgaard & Speer, 1999; Snedeker & Trueswell, 2003; Snow, 1983; Streeter, 1978).

Although the two main aspects of the ambiguous coordinate structure construct, duration and prosodic phrase boundaries by grouping, are well supported prosodic phenomenon, this does not quite elucidate the possible connection between prosodic fluency, prosodic sensitivity and reading comprehension. Our results have shown that participants with average reading comprehension skills produce these durational cues for prosodic phrase boundaries significantly better than those with poorer skills. We have also shown that those with average reading comprehension ability perform better on the listening comprehension task, which utilized the same prosodic and semantic manipulations in the ambiguous coordinate structure, than those with poorer ability. This connection raises the question of what do participants with average reading comprehension skill do better than those with poor reading comprehension skill? Why and how is reading comprehension ability able to predict prosodic fluency and prosodic sensitivity?

Reading comprehension skill is measured usually through silent reading and subsequent silent answering of comprehension questions; at face value, this seems to have little to do with the prosody. However, some researchers would argue this is far from the case. This link between reading comprehension and prosody, crossing both prosodic fluency and prosodic sensitivity, may be connected through the theory of implicit prosody: that when reading silently, the reader creates implicit prosodic representations of the text. The stress, rhythm, F0 contours, pauses, emphasis etc. that are heard in the overt prosody of sentences may be recreated while silently reading. This implicit prosodic representation could theoretically bridge the gap between prosodic production and prosodic sensitivity. If the reader silently generates these prosodic representations, they must be generated accurately (fluency) and they must be attended to successfully (sensitivity), all of course, by the one individual reading.

Implicit prosody can be conceptualized colloquially as “your inner voice when you read”. Janet Fodor, the author of the implicit prosody hypothesis (IPH), found that the differences seen in participants in their parsing preferences of written sentences paralleled the prosodic breaks made whilst talking. Fodor proposed that these dissimilarities between speakers in parsing preference were due an implicit prosodic contour that is present when the participants read the sentences (Fodor, 1998). Depending on what attachment style a participant’s native language favors, the participant’s implicit prosodic breaks may reflect where that participant would pause in their overt prosody of the sentence. These implicit prosodic boundaries are created by the reader in order to parse sentences based on the prosody used in their own speech. Furthering the concept of implicit prosody, in her 2002 paper, Fodor laid out the IPH:

In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction.

This hypothesis postulates that when a reader silently reads a sentence, the implicit prosodic contours created align with the overt prosodic contours heard in speech.

An important aspect of the hypothesis is that the prosody implicitly represented while reading silently mimics the prosody heard in speech; this allows researchers to use stimuli that have produced behavioral data (overt prosody) to attempt to study implicit prosody. Practically applying that reasoning is easier said than done, as researchers cannot directly manipulate implicit prosodic contour. To help address those issues, Fodor (2002) created a list of steps to identify a factor which could be used to study implicit prosody, one of which stated that the factor of interest must have already been documented to affect overt prosody in parsing differences. Additionally, this factor must affect prosodic boundaries and be able to be manipulated in a written sentence; these requirements are important but can pose a challenge for implicit prosody researchers. Although the goal of the current study was not to test for implicit prosody, those steps referenced above give support to using implicit prosody as the link between prosodic fluency, prosodic sensitivity, and reading comprehension. The strength of prior psycholinguistic research on the presence of prosodic phrase boundaries in ambiguous coordinate structures indicates it could reliably be used to manipulate implicit prosody.

Research on implicit prosody usually takes the form of behavioral studies that can utilize eye tracking, or with ERP (event-related potential) methodology. Both types of studies have shown evidence that the same syntactic ambiguities that prosody disambiguates in speech may also be disambiguated by implicit prosody. In other words, implicit prosody can disambiguate syntax. This allows implicit prosody studies to utilize syntactic manipulations that have been shown to have a relationship with prosody, such that prosody aids in disambiguation. Sometimes these manipulations used to elicit an implicit prosodic boundary involve syntactic structure, like garden path sentences, while written sentences may utilize commas as an indicator of a boundary.

Behavioral studies such as eye-tracking and priming methods have been used to explore implicit prosody. One eye tracking study, such as Bader (1998), measured the reading time of ambiguous German sentences to indicate the role of implicit prosody in reading. Participants read ambiguous sentences that either necessitated both a syntactic and a prosodic revision (i.e. garden path sentences) or ambiguous sentences that only required a syntactic revision. Bader found that when both prosodic and syntactic revisions were needed, the reading time significantly increased compared to only a syntactic revision. These results point to the possible presence and utilization of the implicit prosody generated during silent reading, as stimuli which would require prosodic revision in overt prosody showed paralleled delays when reading.

In a reading comprehension experiment, Jun and Bishop (2015) primed participants with unambiguous sentences, which were followed by an ambiguous target sentence. These primes differed only in the presence of a prosodic boundary which affected the attachment of an adjectival clause. Participants answered comprehension questions about the target sentences they silently read. The results showed that when participants were primed with sentences which included a prosodic structure that denoted high attachment, they were significantly more likely to use a high attachment structure to disambiguate the ambiguous target sentence than those primed without a prosodic boundary. These findings support not only the IPH, but also the use of implicit prosodic representations, especially prosodic phrase boundaries, to disambiguate the attachment of a sentence.

ERP methodology, which measures the electrical impulses on the surface of the scalp, provides a temporal map for understanding cognitive processes. ERP studies have been used to study different characteristics of a sentence that may affect various reading processes, such as semantic and lexical processing as well as syntactical processing of reading. These different parts of reading are manifested in ERP experiments through components, or changes in the electrical impulse detected, and these components have been reliably produced by specific manipulations of sentence structure (Swaab, Ledoux, Camblin, & Boudewyn, 2012). The same understanding that certain aspects of reading, when manipulated in written sentences, may elicit different electrical responses has been used to test for the presence of implicit prosodic structures as well.

Several studies have attempted to detect implicit prosodic boundaries with ERP techniques. Researchers have reliably detected distinct shifts in the electrical impulse at locations within a sentence where a prosodic boundary would be placed in overt prosody (Steinhauer & Friederici, 2001; Steinhauer, 2003). These prosodic boundary locations in written sentences would be marked with a comma. Interestingly, in Steinhauer (2003), participants that did not have strict adherence to punctuation did not show that particular electrical shift and performed worse on the comprehension task than those who had more consistent punctuation habits. The discovery of a distinct electrical shift that coincides with the placement of a prosodic boundary presents a strong case for the presence implicit representation of prosodic boundaries in silent reading. It also supports the suggestion that implicit prosody plays a role in the process of silent reading, not to mention a possible link to comprehension.

The previous psycholinguistic literature on overt prosodic boundaries shows its strength in disambiguating syntactically ambiguous stimuli; it is also reliably produced by speakers and used by listeners. Although the psycholinguistic research used similar methodology as our experiments, it alone cannot provide a framework for explaining why participants with normal reading comprehension ability performed better in both experiments compared to those with poor comprehension skills. The IPH could then be a factor that influenced results from the ambiguous coordinate structure constructs. The literature on implicit representations of prosodic phrase boundaries supports both its presence during silent reading, as well as its use in syntactic disambiguation. If implicit prosodic boundaries aid in conveying syntactic information, then implicit prosodic structures may directly affect the comprehension of sentences. It is this role that may tie implicit prosody to reading comprehension ability. Prosodic structure may be generated implicitly during silent reading as a mechanism to facilitate comprehension. This could mean that participants who display strong reading comprehension skills may also be skilled in generating implicit prosody. Furthermore, the stimuli used in the ambiguous coordinate structure construct were disambiguated by prosody, so this would place greater importance on such prosodic skill in successful comprehension. This, however, begs the question what does it mean to be skilled in generating implicit prosody?

Successfully creating implicit prosody may contain the elements which our two experiments tested for: accurate production of important prosodic characteristics, and the ability to detect and utilize those same characteristics. In order to generate implicit prosody while reading, the reader has to silently produce those characteristics as if reading aloud, as implicit prosody should resemble overt prosody (Fodor, 2002). Then, the reader must identify the prosodic features generated, and be able to understand the information it carries; which parallels what was measured in the prosodic sensitivity experiment. In other words, good implicit prosodic generation requires good prosodic fluency *and* sensitivity abilities.

The prosody of written sentences is not always easily mapped. Therefore learning to map that prosody is challenging but important, especially as prosodic cues are used to chunk sentences to help comprehension (Schreiber, 1991). Those readers who learn how to successfully recognize and produce that prosody should also have increased comprehension ability, while those who do not learn prosodic structure as well could be less likely to improve their reading ability. This supports the idea of implicit prosody as a possible moderator variable for our results, in that those who have poor prosodic skills (as evidenced by our results) may also generate poor implicit prosody, which relates to poor reading comprehension skill.

Although the idea of implicit prosody as the connection between prosodic fluency, prosodic sensitivity, and reading comprehension ability has some conceptual support, our studies were not without some limitations. Some of our acoustic measurements are not incredibly precise. The prosodic feature of F0 change is considered a noisy feature; it can be hard to accurately capture F0 differences in participants. Our measurements of duration as lengthening of a boundary adjacent word plus following silence are coarse measurements as well. With additional measurements, such as only analyzing the length of the final syllable of boundary adjacent words or further utilizing ToBI annotation within our imitation analysis, our data would be cleaner and better reflect the prosody of our participants. An inclusion of a perceptual measure into the imitation task would strengthen our prosodic fluency analysis, as the imitations participants produce is based on their perceptions of the model sentences. A perceptual measure would ensure the validity of the salience of the prosodic cues in those model sentences.

Other limitations pertained to our participants and possible confounds. Our correlation analysis testing the results between our two studies did not find significance, but this may be due to a lack of power. The correlation analysis only contained 18 participants who had completed both experiments for us to run a within-subjects analysis, so for future work a larger sample may be necessary. The significant differences between the participants in the listening comprehension experiment was also concerning; ensuring that only IQ and KTEA score differed significantly between control and poor comprehenders may have improved results. Additionally, certain constructs spanning the two experiments (focus not phrase and scope ambiguities) were found to have not reliably measured or produced the prosodic characteristics desired. The prosody and sentence structures we did analyze were more limited because of these exclusions. Another possible issue is that the difference in performance between the groups was due to a difference in another skill. Although IQ and age were controlled for, those skills or another may have affected the results of this study. Although we do suggest that implicit prosody plays a role in modulating the relationship of prosody and reading comprehension, the conclusions are limited as we did not set out to measure implicit prosody.

Future directions from these experiments may include a possible ERP study using the ambiguous coordinate stimuli, perhaps in the form of the prosodic sensitivity paradigm (as a production task would interfere with data collection). Testing with ERP methodology could be a way to further support the inclusion of the IPH within the relationship between prosody and reading comprehension. In considering further studies, integrating the expectations of imitation tasks within the motivations of the research will be important. Understanding the differences between an imitation task and a production task, along with including a production task, will strengthen further research. Additionally, analyzing narrative stimuli (to build upon our current repetition stimuli) for a production task would add an interesting layer, in that discourse prosody may manifest differently and give new insight into prosodic fluency and reading comprehension. Integrating an eye-tracking component with the prosodic sensitivity paradigm would give detail into when in the stimuli participants looked at what image on the screen; pinpointing their gaze would allow us to better understand if the prosodic cue given was accurately utilized by the participant. Our current study only utilized two comprehension groups, a control and a poor ability group. In future experiments, adding a group of participants who display exceptional reading comprehension ability could further elucidate prosodic differences with regards to comprehension skill.

This study, with its two experiments exploring prosodic fluency and prosodic sensitivity, elicited some interesting findings about the relationship between reading comprehension and prosody. Prior literature on prosody and reading comprehension did not adequately encapsulate a definition on prosody and reading fluency, and so our imitation experiment sought to define one. Our definition posits that prosodic fluency is the ability to appropriately produce prosodic characteristics as determined by the syntactic and semantic context. Although most of the constructs tested proved to not render significant results, the ambiguous coordinate structure elicited significant differences between reading comprehension levels in both experiments. By assessing both prosodic fluency and sensitivity, our study encompasses prosody more fully than previous literature. Our findings have implications that are greater than just prosodic fluency or prosodic sensitivity and indicate there is relationship between those two aspects of prosody that connects with reading comprehension ability. They may shed light on the implicit aspect of prosody that has not been widely discussed as relating to reading comprehension. Moving forward, perhaps implicit prosody will be considered as a piece to the puzzle regarding prosody and reading comprehension.

References

Allington, R. L.. (1983). Fluency: The Neglected Reading Goal. *The Reading Teacher*, *36*(6), 556–561.

Baayen, R. H. (2008). Analyzing linguistic data: A practical introduction to statistics using R. Cambridge University Press.

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In *Reanalysis in sentence processing* (pp. 1-46). Springer Netherlands.

Benjamin, R. G., & Schwanenflugel, P. J. (2010). Text complexity and oral reading prosody in young readers. *Reading Research Quarterly*, *45*(4), 388-404.

Binder, K. S., Tighe, E., Jiang, Y., Kaftanski, K., Qi, C., & Ardoin, S. P. (2013). Reading expressively and understanding thoroughly: An examination of prosody in adults with low literacy skills. *Reading and writing*, *26*(5), 665-680.

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer [Computer program]. Version 5.3. 03. 2011.

Breen, Kaswer, Van Dyke, Krivokapić, and Landi. (2015) *Prosodic fluency predicts reading comprehension ability*. Manuscript submitted for publication.

Breen, M., Watson, D. G., & Gibson, E. (2011). Intonational phrasing is constrained by meaning, not balance. *Language and Cognitive Processes*,*26*(10), 1532-1562.

Carlisle, J. F. (1988). Knowledge of derivational morphology and spelling ability in fourth, sixth, and eighth graders. *Applied Psycholinguistics*, *9*(03), 247-266.

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

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

Carlson, K., Frazier, L., & Clifton, C. (2009). How prosody constrains comprehension: A limited effect of prosodic packaging. *Lingua*, *119*(7), 1066-1082.

Clin, E., Wade-Woolley, L., & Heggie, L. (2009). Prosodic sensitivity and morphological awareness in children's reading. *Journal of Experimental Child Psychology*, *104*(2), 197-213.

Cole, J., & Shattuck-Hufnagel, S. (2011). The phonology and phonetics of perceived prosody: What do listeners imitate?. In *Interspeech* (pp. 969-972).

Cooper, R.M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6* (1974), pp. 84–107

Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of the Late Closure strategy in Spanish. *Cognition*,*30*(1), 73-105.

Cutler, A., Oahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40,* 141–201.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*(2), 292-314.

Dankovicová, J., Pigott, K., Wells, B., & Peppé, S. (2004). Temporal markers of prosodic boundaries in children's speech production. *Journal of the International Phonetic Association*, *34*(01), 17-36.

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

Dowhower, S. L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory Into Practice*, *30*(3), 165-175.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America*, *80*(2), 402-415.

Farnetani, E., Torsello, C. T., & Cosi, P. (1988). English compound versus non-compound noun phrases in discourse: An acoustic and perceptual study. *Language and speech*, *31*(2), 157-180.

Féry, C., & Krifka, M. (2009). Information structure: Notional Distinctions, Ways of Expression. Pennsylvania State University.

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, *36*(4), 680-703.

Fodor, J. D. (1998). Learning to parse?. *Journal of psycholinguistic research*,*27*(2), 285-319.

Fodor, J. D. (2002). Psycholinguistics cannot escape prosody. In *Speech Prosody 2002, International Conference*.

Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244-249.

Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of memory and language*, *63*(3), 367-386.

Gorman, K., Howell, J., & Wagner, M. (2011). Prosodylab-aligner: A tool for forced alignment of laboratory speech. *Canadian Acoustics*, *39*(3), 192-193.

Hintze, J. M., Ryan, A. L., & Stoner, G. (2003). Concurrent validity and diagnostic accuracy of the dynamic indicators of basic early literacy skills and the comprehensive test of phonological processing. *School Psychology Review*, *32*(4), 541-556.

Holliman, A. J., Wood, C., & Sheehy, K. (2008). Sensitivity to speech rhythm explains individual differences in reading ability independently of phonological awareness. *British Journal of Developmental Psychology*,*26*(3), 357-367.

Holliman, A. J., Williams, G. J., Mundy, I. R., Wood, C., Hart, L., & Waldron, S. (2014). Beginning to disentangle the prosody-literacy relationship: A multi-component measure of prosodic sensitivity. *Reading and Writing*, *27*(2), 255-266.

Hudson, R. F., Pullen, P. C., Lane, H. B., & Torgesen, J. K. (2009). The complex nature of reading fluency: A multidimensional view. *Reading & Writing Quarterly*, *25*(1), 4-32.

Huettig, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation.*Acta psychologica*, *137*(2), 151-171.

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, *58*(2), 541-573.

Jansen, W., Gregory, M. L., & Brenier, J. M. (2001). Prosodic correlates of directly reported speech: Evidence from conversational speech. In *ISCA tutorial and research workshop (ITRW) on prosody in speech recognition and understanding*.

Jun, S. A., & Bishop, J. (2015). Priming implicit prosody: Prosodic boundaries and individual differences. *Language and Speech*, 0023830914563368.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test*. John Wiley & Sons, Inc..

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity.*Journal of Memory and Language*, *40*(2), 153-194.

Kuhn, M. R., Schwanenflugel, P. J., Meisinger, E. B., Levy, B., & Rasinski, T. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, *45*(2), 230-251.

Kurumada, C., Brown, M., & Tanenhaus, M. K. (2012). Pragmatic interpretation of speech: It looks like speech adaptation. In *Proceedings of the 35th annual meeting of the Cognitive Science Society. Sapporo, Japan*.

Kutik, E. J., Cooper, W. E., & Boyce, S. (1983). Declination of fundamental frequency in speakers' production of parenthetical and main clauses. *The Journal of the Acoustical Society of America*, *73*(5), 1731-1738.

Lai, S. A., George Benjamin, R., Schwanenflugel, P. J., & Kuhn, M. R. (2014). The longitudinal relationship between reading fluency and reading comprehension skills in second-grade children. *Reading & Writing Quarterly*,*30*(2), 116-138.

Lehiste, I. (1970). Suprasegmentals. Cambridge, Massachusetts: MIT Press.

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.

Miller, J., & Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly*, *43*(4), 336-354.

Nooteboom, S. (1997). The prosody of speech: melody and rhythm. *The handbook of phonetic sciences*, *5*, 640-673.

Paige, D. D., Rasinski, T., Magpuri-Lavell, T., & Smith, G. S. (2014). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research*, *46*(2), 123-156.

Pardo, J. S. (2013). Measuring phonetic convergence in speech production.*Frontiers in psychology*, *4*(559), 183-195.

Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* (Doctoral dissertation, Massachusetts Institute of Technology).

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. *Intentions in communication*, *271*, 311.

Pinnell, G. S., Pikulski, J. J., Wixson, K. K., Campbell, J. R., Gough, P. B., & Beatty, A. S. (1995). Listening to children read aloud. Washington, DC: Office of Educational

Plag, I., Kunter, G., Lappe, S., & Braun, M. (2008). The role of semantics, argument structure, and lexicalization in compound stress assignment in English. *Language*, *84*(4), 760-794.

Price, Ostendorf, Shattuck-Hufnagel, & Fong. (1991). The use of prosody in syntactic disambiguation. *The Journal of the Acoustical Society of America*, *90*, 2956–2970

Pruitt, K., & Roelofsen, F. (2014). The interpretation of prosody in disjunctive questions. *Linguistic inquiry*, *44*(4), 632-650.

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

Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics*, *1*(1), 75-116.

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

Schwanenflugel, P. J., Hamilton, A. M., Kuhn, M. R., Wisenbaker, J. M., & Stahl, S. A. (2004). Becoming a Fluent Reader: Reading Skill and Prosodic Features in the Oral Reading of Young Readers. *Journal of educational psychology*, *96*(1), 119.

Silverman, K. E., Beckman, M. E., Pitrelli, J. F., Ostendorf, M., Wightman, C. W., Price, P., ... & Hirschberg, J. (1992, October). TOBI: a standard for labeling English prosody. In *ICSLP* (Vol. 2, pp. 867-870).

Silverman, S. W., & Ratner, N. B. (1997). Syntactic complexity, fluency, and accuracy of sentence imitation in adolescents. *Journal of Speech, Language, and Hearing Research*, *40*(1), 95-106.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and language*, *48*(1), 103-130.

Snow, C. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review*, *53*(2), 165-189.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and language*, *86*(1), 142-164.

Steinhauer, K., & Friederici, A. D. (2001). Prosodic boundaries, comma rules, and brain responses: The closure positive shift in ERPs as a universal marker for prosodic phrasing in listeners and readers. *Journal of Psycholinguistic Research*, *30*(3), 267-295.

Streeter, L. A. (1978). Acoustic determinants of phrase boundary perception.*The Journal of the Acoustical Society of America*, *64*(6), 1582-1592.

Swaab, T. Y., Ledoux, K., Camblin, C. C., & Boudewyn, M. A. (2012). Language-related ERP components. *Oxford handbook of event-related potential components*, 397-440.

Swerts, M., & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and speech*, *37*(1), 21-43.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634.

Tighe, E. L., & Schatschneider, C. (2014). Examining the Relationships of Component Reading Skills to Reading Comprehension in Struggling Adult Readers A Meta-Analysis. *Journal of learning disabilities*, 0022219414555415.

Truckenbrodt, H. (2006). Phrasal stress. *The encyclopedia of languages and linguistics*, *9*, 572-579.

Vitz, P. C. (1972). Preference for tones as a function of frequency (hertz) and intensity (decibels). *Perception & Psychophysics*, *11*(1), 84-88.

Wagner, M. (2005). *Prosody and recursion* (Doctoral dissertation, Massachusetts Institute of Technology).

Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and cognitive processes*, *25*(7-9), 905-945.

Walczyk, J. J., Marsiglia, C. S., Johns, A. K., & Bryan, K. S. (2004). Children's compensations for poorly automated reading skills. *Discourse Processes*, *37*(1), 47-66.

Wells, B., Peppé, S., & Goulandris, N. (2004). Intonation development from five to thirteen. *Journal of Child Language*, *31*(04), 749-778.

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. Psychological Corporation.

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

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries.*The Journal of the Acoustical Society of America*, *91*(3), 1707-1717.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside Publishing.

Worthington, C. F. (1987). Taking the test: Kaufman Test of Educational Achievement, Comprehensive Form and Brief Form. *Journal of Counseling and Development*, *65*, 325–327.

Zutell, J., & Rasinski, T. V. (1991). Training teachers to attend to their students' oral reading fluency. *Theory Into Practice*, *30*(3), 211-217.

Appendix A

List of stimuli used in Experiment 1

Declarative Statements

1. Emily has painted a melon.
2. Anna will now play the piano.
3. Allison is starting ninth grade.
4. Stewart will go to college next year.
5. Alex is waiting at the bar.
6. Nicole finished making dinner.
7. Gwen gave a presentation on Monday.
8. Annie is buying an orange.
9. Dan is looking at an orange.
10. Ann is going to take the money.

Yes-No Questions

1. Has Emily painted a melon?
2. Will Anna now play the piano?
3. Is Allison starting ninth grade?
4. Will Stewart go to college next year?
5. Is Alex waiting at the bar?
6. Has Nicole finished making dinner?
7. Did Gwen give a presentation on Monday?
8. Is Annie buying an orange?
9. Is Dan looking at an orange?
10. Is Ann going to take the money?

Basic Quotatives

1. "That sounds wonderful!" said Jane.
2. "I like it." said Andrea.
3. "Good for you!" said Mark.
4. "It's not bad." said Leslie.
5. "I want one!" Evan screamed.
6. "How awful!" Johnny said.
7. "It's for you." Adam said.
8. "That's not good." Peter replied.
9. "Let's play!" exclaimed Luther.
10. "That's new!" Ronda said.

Ambiguous Coordinates (Two-One)

1. Ann and Bobby, or Nancy, will come.
2. Ann will teach Abe and Bob, or Lenny.
3. John will paint Al and Dan, or Nina.
4. George and Donald, or Larry, will take art.
5. Sal studies with Ally and Nina, or Louie.

6. Mathew carpools with Jim and Zackary, or David.
7. Dan and Ryan, or Nicole, had dinner.
8. Ann and Doug, or Lee, will take you home.
9. Dan and Lenny, or Edward, will leave.
10. Ann and Roy, or Dan, will leave London.

Ambiguous Coordinates (Two-Two)

1. Ann, and Bobby or Nancy, will come.
2. Ann will teach Abe, and Bob or Lenny.
3. John will paint Al, and Dan or Nina.
4. George, and Donald or Larry, will take art.
5. Sal studies with Ally, and Nina or Louie.
6. Mathew carpools with Jim, and Zackary or David.
7. Dan, and Ryan or Nicole, had dinner.
8. Ann, and Doug or Lee, will take you home.
9. Dan, and Lenny or Edward, will leave.
10. Ann, and Roy or Dan, will leave London.

Relative Clauses

1. The blue motel, on the corner of Hollywood and Vine, is haunted.
2. Their father, who was in the army, stood in the yard.
3. After the rain, which ended at nine, we walked around the pond.
4. Every Monday, at the crack of dawn, Andy gets up to milk the cows.
5. Now, after many years, lettuce is in demand again.
6. Laurie's grandma, who is a lawyer, did not like the brochure.
7. The girl, who was injured, ran from the tornado.
8. The Dude, who was on a mission, broke into the house.
9. That evening, while cooking, Paula knew something was wrong.
10. The driveway, which was muddy, led to the mansion.
11. The novel, which she finished, sat on the nightstand.
12. The star player, who joined the team, was a receiver.
13. The widow, who was devastated, received many flowers.
14. The journey, which I took alone, felt like it would never end.
15. The groom, who was on edge, drank too much.
16. The chaperones, who volunteered, forgot to come.
17. The apple, which she bought at the farm, was fresh.
18. The room, which had a red chair, caught Mandy’s eye.
19. His wagon, which only had three wheels, was noisy.
20. The play, which Bobby wrote junior year, was about zombies.

Unambiguous Coordinates

1. Ann has a dog, a pen, and a mug.
2. Ann bought a rose, a bird, and jelly.
3. John wants a bagel, butter, and jam.
4. George loves history, math, and English class.
5. Alex takes a car, a taxi, or a train to work.

6. Nicole made chicken, bread, and pasta.
7. Ben talked about the baby, the doll, and the lion.
8. Ann bought a pie, a bag, and a lime.
9. Dan plays with Rob, Nathan, and Morgan.
10. Ann looked at Ed, Mary, and Nellie.

Appendix B

List of stimuli used in Experiment 2

Accent Disambiguation

1. Sam created some of his best work in a **darkroom vs. dark room**
2. Joe parked his car outside **the greenhouse vs. green house**
3. Nancy passed by **a hotdog vs. hot dog** on the sidewalk in the park.
4. Tommy used a **blackboard vs. black board** to accent the wall.
5. Albert grabbed the **tophat vs. top hat** off the shelf.
6. John walks by the **White House vs. white house** every day on his way to school.
7. Jenna saw a **goldfish vs. gold fish** in the store window.
8. Shannon pushed the **highchair vs. high chair** closer to the table.
9. His hometown is known for its distinct type of **bluegrass vs. blue grass**.
10. The **tightrope vs. tight rope** was so thin you could barely see it in the light.
11. Sally tossed a **softball vs. soft ball** to her little sister.
12. Harold was eager to get his **blackbelt** vs **black belt** after school.
13. Gerald used **redwood** vs. **red wood** to make a bookcase.
14. In the gift shop, Gina bumped into a **redcoat** vs. **red coat** on display.

Subject Focus Object Focus

1. The rabbit will eat carrots.
2. John likes to play soccer.
3. The apple will go in the basket.
4. Sarah will order a pizza.
5. The book is on the shelf.
6. John is drinking coffee.
7. The muffin contains cranberries.
8. Jack will drive to California.
9. Amy went to the bookstore.
10. The girl skipped in the garden.
11. Beth washed the car.
12. Ben fixed the television.
13. The cat sits by the window.
14. The bikers always wear helmets.

Contour Ambiguity

1. It looks like a backpack.
2. It looks like a bathtub.
3. It looks like a beaver.
4. It looks like a cellphone.
5. It looks like a chessboard.
6. It looks like a hammer.
7. It looks like a zebra.
8. It looks like a lemon.

9. It looks like a necklace.
10. It looks like a panda.
11. It looks like a penguin.
12. It looks like a puppy.
13. It looks like a bunny.
14. It looks like a camel.
15. It looks like a kitty.

Disjunctive Questions

1. Does Paula sing or dance?
2. Does Roger plan to mow the grass or take out the recycling?
3. Is Pamela going to knit a scarf or buy a sweater?
4. Was Samantha going to walk the dog or feed the cat?
5. Did Phil use sunscreen or wear a hat?
6. Is Bruce going to show us a map or draw us a picture?
7. Did William attend the meeting or send an e-mail?
8. Did Sally bring wine or make dessert?
9. Would Emily like to visit the aquarium or go to the zoo?
10. Did Eddie spend all night watching movies or playing video games?
11. Was Pat going to wash the dishes or mop the floor?
12. Do the kids need a snack or a bathroom break?
13. Did Alan write her a poem or buy her flowers?
14. Does Petra want to get drinks or go dancing?

Ambiguous Phrasing

1. Tap the cow with the lollipop.
2. Lift the horse with the towel.
3. Pinch the bear with the barrette.
4. Scratch the dog with the coin.
5. Pet the bear with the pom pom.
6. Tickle the dog with the pencil.
7. Spin the cow with the spoon.
8. Touch the frog with the flower.
9. Rub the elephant with the egg.
10. Poke the giraffe with the pen.
11. Move the leopard with the blanket.
12. Pick up the mouse with the shirt.
13. Whack the snake with the toy.
14. Scratch the lion with the branch.

Coordination

1. Sarah and Will or Molly paid for dinner
2. Alan and Paul or Lizzy can go with Grandpa
3. Bill and Julie or Elaine were in an accident
4. Jacob and Kelly or Fran took the baby to the park
5. Barry and Susan or Earl cleaned the kitchen

6. Connie and Martin or Dave will attend the conference
7. Henry and Lauren or Zoe washed the dishes
8. Roger and Sal or Mary walked the dog
9. Bob and Jackie or Nora got food poisoning
10. George and Abe or Wendy will go to the game
11. Bernie and Meghan or Lisa planned the party
12. Mel and Dan or Lenny did the project
13. Jill and Larry or Dillon were stuck in traffic
14. Ben and Michael or Lilly raked the leaves
15. Joe and Annie or John watched a movie