THE EFFECTS OF DISTAL PROSODIC CUES ON SPEECH PERCEPTION IN ADULTS WHO STUTTER

By

Kaitlyn Marie Ayres

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Communicative Sciences and Disorders - Master of Arts

2015

ABSTRACT

THE EFFECTS OF DISTAL PROSODIC CUES ON SPEECH PERCEPTION IN ADULTS WHO STUTTER

By

Kaitlyn Marie Ayres

Stuttering is a neurodevelopmental disorder characterized by the disruption of rhythmic flow and timing in speech production. Speech productions of people who stutter are influenced by the perception of prosodic aspects of their own speech, suggesting that people who stutter may have difficulty in perceiving prosodic elements of their own speech causing disfluency. To investigate this possibility, adults who stutter (n=15) and adult fluent speakers (n=13) were compared in their ability to perceive distal rhythmic and timing cues in speech. Adults who stutter exhibited perception of the distal prosodic cues and an ability to internally generate rhythm and timing for parsing of syllables and words in speech. These findings provide evidence of intact perception of rhythm and timing of speech in adults who stutter.

ACKNOWLEGMENTS

I would like to thank my thesis committee members, Dr. Laura Dilley, Dr. Soo-Eun Chang, and Dr. Amanda Hampton Wray for all of their encouragement, support, and guidance with this research. Additionally, I thank Elizabeth Wieland, Gregory Spray, Kristin Hicks, Saralyn Rubsam, Zachary Ireland, and other members of the MSU Speech Perception-Production Laboratory and MSU Speech-Neurophysiology Laboratory for their generous assistance with this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

Children who stutter …………………………………………………………………………………CWS

Persons who stutter ……………………………………………………………………………...…..PWS

Delayed auditory feedback……………………………………………………………………….…..DAF

Basal ganglia thalamocortical network…………………………………………………….....BGTC

Introduction

Stuttering is a speech disorder marked by frequent hesitations, prolongations and/or repetitions of sounds, syllables and words. These types of stuttering disfluencies characterize the disorder as a disruption of the rhythmic flow of speech production (World Health Organization, 2010). Although many theoretical models suggest possible underlying mechanisms or causes of stuttering, none have thus proven to provide a completely sufficient explanation (e.g., Packman, Code, & Onslow, 2007).

All components of language production, including pitch and rhythm characteristics (referred to as prosody), have been a recurring focus in this search for a cause of stuttering. Various previous research has shown that aspects of prosody play important roles in the production and perception of fluent speech (Bergmann, 1986; Besozzi & Adams, 1969; Hall, Amir, & Yairi, 1999; Packman et al., 2007; Packman, Onslow, Richard, & Van Doorn, 1996; Wingate, 1966). The known symptoms of dysrhythmic flow in the motor speech production of persons who stutter (PWS) have led to proposals that an underlying deficit in PWS may be associated with the prosodic aspects of rhythm and timing used in speech (Alm, 2004). Stuttering has been compared to other disorders such as dystonia or Parkinson's disease, due to the similarity in symptoms of disrupted rhythm or timing in motor movements. These disorders are classified as basal ganglia motor disorders, and thus it has been suggested that an impaired ability of the basal ganglia in generating timing cues for initiating motor speech is a core dysfunction in stuttering (Alm, 2004).

Despite not having revealed the underlying deficit of stuttering, research has identified various conditions in which stuttering is decreased, most of which utilize aspects of rhythm and timing. Studies have shown that speech productions of people who stutter (PWS) can be influenced by the perception of prosodic aspects of their own speech (Hargrave, Kalinowski, Stuart, Armson, & Jones, 1994; Wingate, 1966). Packman and colleagues (1996) theorized that the beneficial effects of rhythmic speech on stuttering are due to the reduced variability of syllabic stress provided by the external rhythm (i.e., the V Model). The patterns of syllabic stress are influenced by the rhythmic patterns of speech perception in PWS. Kent (1984) suggested that the difference in ability to generate rhythmic patterns for speech production is a principal distinction between people who stutter and fluent speakers. These findings throughout various research have been supported by conditions in which fluency is enhanced, including choral reading (Adams & Ramig, 1980; Ingham & Carroll, 1977), singing (Glover, Kalinowski, Rastatter, & Stuart, 1996), and speaking in time with a metronome (Wingate & Howell, 2002). These evidence-based, fluency-inducing conditions all share one key aspect: the provision of an external rhythm or external pacing of speech. The fluency-inducing effects of an external pacing signal suggest PWS may have an impaired internal pacing signal. Knowing that elements of speech productions are generated from the stored perceptions of speech (Guenther, 1994), we can infer that if PWS exhibit a deficit in perceiving rhythmic prosodic cues in speech, it may be affecting their ability to internally generate these prosodic cues in their speech productions.

In this study we investigated the possibility of PWS exhibiting an impaired utilization of prosodic elements in the perception of their own speech, relating to their disfluency in speech production. Any deficits found in the perceptions of prosodic characteristics of speech (pitch rhythm and timing) by PWS support the idea that disruption of internal rhythm generation from perception of rhythm and timing in their own speech could be a contributing factor to the core deficits of stuttering. The existence of deficits in the rhythmic-prosodic components of speech for PWS are supported by literature which shows treatments involving external pacing help improve their fluency of speech (Ingham & Carroll, 1977; Wingate & Howell, 2002; Glover, Kalinowski, Rastatter, & Stuart, 1996). If PWS exhibit a deficit in perceiving these rhythmic prosodic cues in speech, it may be affecting the ability to internally generate rhythmic pacing of speech, which would be evident in both the perception and production of speech.

Recent findings from functional and structural magnetic resonance imaging studies have shown neurological evidence of these possible deficits in internal beat generation that are used for the rhythmic production of speech (Chang & Zhu, 2013; Grahn & McAuley, 2009;Toyomura et al., 2015). Rhythm processing and internal generation of a beat are neurologically supported by the basal ganglia thalamocortical (BGTC) network, which involves connectivity among the basal ganglia, supplementary motor area (SMA), and pre-motor and auditory regions (Grahn & McAuley, 2009). These cortical and sub-cortical regions are commonly activated in tasks that involve motor timing (such as speech) and are strongly associated with temporal perceptions and processing of auditory stimuli

(Harrington et al., 1998b; Schubotz et al., 2000; Rao et al., 2001; Nenadic et al., 2003; Coull, 2004; Grahn and Brett, 2007; Hazeltine et al., 1997; Harrington and Haaland, 1999). Grahn & McAuley (2009) support these notions by finding greater activity of the pre-motor and supplementary motor cortices and basal ganglia in adults while making timing judgments in an auditory perception task. Additionally, Chang and Zhu (2013) showed that children who stutter have attenuated connectivity within the BGTC network compared to age-matched controls. Toyomura et al. (2015) suggested that the basal ganglia is involved in the improvement of speech fluency by utilizing the perceptions of an externally triggered rhythm. A deficit in rhythm perception was found in children who stutter (CWS) by Wieland, McAuley, Dilley, and Chang (2015), also providing evidence for a deficit in speech perceptions of PWS. Differences between internally generated pacing for speech and externally-triggered pacing may be important for understanding the underlying neuromechanisms of disfluency in persons who stutter based on the fact that stuttering occurs in self-paced speech but is temporarily decreased with the presence of an external pacing (Toyomura, 2015).

All of these findings provide increasing evidence for a possible deficit in the BGTC network in PWS that could potentially be affecting the ability to internally generate rhythm and timing for motor production. These results also suggest the correlation between a possible deficit in the internal rhythm generation within the self-perceptions of speech by people who stutter with concomitant effects on the fluency of speech production.

Although previous research has alluded to the possibility of PWS exhibiting deficits in prosodic elements, none have before looked at these effects on the speech perception of PWS. Further, how the perception of these prosodic cues affects linguistic function in segmentation and syllable parsing has never been investigated in PWS. In the past 35 years, research focusing on the influence of prosody on speech perception targeted how prosodic properties of local pitch and timing (proximal to a word itself) affected the segmentation and rhythmic properties of that word (Cutler, Dahan, van Donselaar, 1997). Prosodic characteristics of spoken language influencing the processing and parsing of speech have been shown in control populations across many different languages. The literature has shown that listeners tend to perceive stressed syllables (i.e., syllables that include high pitch and longer duration) as the beginning of new words and likewise, syllable lengthening as signaling the end of a word or phrase (Cutler, Dahan, & van Donselaar, 1997; Salverda et al., 2007. While research on the effects of prosody on the parsing of speech began with investigation of proximal (local) prosodic characteristics, research has evolved into including investigation of distal (non-local) prosody on the parsing of speech (Cutler, Dahan, van Donselaar, 1997; Dilley & McAuley, 2008). Valid and reliable ways to gather measures of how people perceive particular elements of distal (non-adjacent) prosodic cues has been established and tested with typical control populations (Dilley & McAuley, 2008; Dilley & Pitt, 2010; Morrill, Dilley, and McAuley, 2014). Results from these studies concluded that typical-developing adults utilize distal (i.e., distant, non-adjacent) prosodic cues of rhythm and timing of speech in their parsing of syllables and lexical segmentation of words. This was demonstrated by the fact that alterations in distal prosodic cues yielded significantly different speech perceptions for listeners of proximal material (Dilley and McAuley, 2008; Dilley and Pitt, 2010; Morrill et al., 2014).

Dilley and Pitt (2010) demonstrated clear evidence that a manipulation of distal speech rate induced typically-developing listeners to segment words differently. The speech perception task utilized in this study involved judgments on sentences that included a target syllable with a function word (e.g., *or* in a sentence like *Deena didn't have any leisure or time*), presented auditorily. When the distal speech rate was at a normal rate, listeners reported hearing the function word. However, when the distal speech rate was slowed, listeners did not perceive the function word. Despite the fact that the acoustic properties of the proximal, targeted syllable (function word) of each sentence being acoustically identical, the perceptions of the target syllable were altered based on the manipulated rate of the distal context. This finding demonstrates that listeners from a typically developing control population are sensitive to the distal speech rate. These manipulations also demonstrated the function of prosodic cues in speech on the lexical segmentation of speech perception in typical-developing adults.

Other research has revealed additional sensitivities in typically-developing listeners to aspects of distal prosody. For example, a paradigm developed by Dilley and McAuley (2008) and extended by Morrill, Dilley, and McAuley (2014) revealed effects of distal rhythmic cues on linguistic parsing of syllables into word units. In these studies, participants heard auditory speech sequences containing two disyllabic, non-compound words continued with a string of four syllables (e.g. banker helpful [tɑɪ mɝ dɚ bi]). The string of final four syllables had ambiguous lexical structure and therefore could be organized more than one way (e.g., *timer derby, tie murder bee*). The organization of the final four syllables was shown in these studies to be dependent on the fundamental frequency (F0) and duration of the distal context. These studies used sentences such as banker helpful [tɑɪ mɝ dɚ bi], in which the lexical organization of the final four syllables is ambiguous (*timer derby* or *tie murder bee*). The experiments utilized two distinct distal prosodic conditions that were expected to lead to perceptual grouping of syllables in two different ways:. In one distal prosodic pattern, a repeated low (L)-high (H) alternation of F0 on distal syllables was predicted to lead to grouping of L and H pitch elements across the entire sequence, which was predicted to generate perception of a disyllabic word (e.g., *derby*); these were referred to as Disyllabic contexts. In another distal prosodic pattern, a repeated high-low alternation of F0 was predicted to lead to grouping of H and L pitch elements across the entire sequence, which was predicted to generate perception of a monosyllabic final word (e.g., *bee*); these were referred to as Monosyllabic contexts. In more detail, for both the monosyllabic and disyllabic contexts, the final three syllables of each target sequence received a high (H), low (L), high (H) F0 pattern (one F0 target per syllable). The first five syllables of each target sequence varied in F0 pattern. The first five syllables of the disyllabic context items received a L1-H2-L3-H4-L5-pattern, with one F0 target, H or L, on each syllable. The monosyllabic pattern received a H1-L2-H3-L4- HL5- pattern with one F0 target for each of the first four syllables, and a fall in F0 from H to L on the fifth syllable. These distal prosodic context manipulations predicted the perceptual grouping of the final syllables in the target sequence, in which the Disyllabic context caused a relatively stronger prosodic boundary to be heard before the penultimate syllable (e.g., der) than before the final syllable (e.g., bee), so that listeners would group the final two syllables as a single disyllabic word (e.g., derby). In contrast, the Monosyllabic context caused the reverse with a stronger prosodic boundary to be heard before the final syllable (e.g., bee) than before the penultimate syllable, so that listeners would hear a final monosyllabic word (e.g., bee). The Disyllabic context yielded a (L1-H2-)(L3-H4-)(L5-H6-)(L7-H8) grouping of the final three syllable sequence, displaying a larger boundary before syllable 7 than syllable 8, thereby critically yielding a disyllabic final word report (e.g., derby). In contrast, the Monosyllabic context yielded a (H1-L2-)(H3-L4-)(HL5-) (H6-L7-)(H8...) grouping of the final three syllables, displaying a larger boundary before syllable 8 than before syllable 7. -The pattern of responses by typical listeners was as predicted: there was a strong tendency for listeners to hear the same acoustic material as ending in disyllabic words in Disyllabic contexts but as ending in monosyllabic words in Monosyllabic contexts. Results for these studies showed that typical-developing adults perceived distal prosodic cues of intonation rhythm in speech and, that their perceptions manipulated their parsing of syllables in speech perception.

The studies reported above have provided reliable experiment paradigms for measuring effects of distal prosodic cues in typical adults. The goal of this thesis is to extend these paradigms to examine speech perceptions in PWS. These paradigms require listeners to not only perceive the distal prosodic cues but also, crucially, to internally generate rhythm and timing from the externally prompted distal prosodic cues in order to parse the proximal speech material in the predicted way.

Based on previous literature, it seems plausible that PWS may have a deficit in their ability to internally generate intonation rhythm and timing used for speech production. This is supported by the fact that external rhythms, rates, and beats are evidence-based to help with fluency in this clinical population (Packman et al., 2007). Studies of the effects of intonation, rhythm, and duration on word perceptions have never before been conducted with PWS. These studies by Dilley and Pitt (2010) and Morrill, Dilley, and McAuley (2014) provide an effective way to measure the influence of prosody in speech perception and allow measuring sensitivity to prosody by indexing linguistic perception. In the present study, the sensitivity to prosodic cues was investigated in both a control and clinical population; this comparison allowed for assessing any group differences and provided insight into possible perceptual gaps in PWS. Paradigms utilizing manipulations of distal prosodic cues were utilized due to the demands required by the listener for distal manipulations compared to proximal manipulations. When the manipulations of prosodic cues are distal to the lexically ambiguous target portion of speech, there is a requirement for internally generating the rhythm and timing signaled distally which prior results suggest should affect perception of speech. The carry-over of this internally generated rhythm is a valuable measure for investigation of speech perception within PWS, one that could not be tested using experimental paradigms that involved proximal prosodic cues.

This study is motivated by the question of whether people who stutter (PWS) show a difference in the effects of distal prosodic cues, (pitch rhythms and rate of speech) on speech perceptions. This question will be investigated from the results of two experimental paradigms, each utilizing manipulations to different prosodic characteristics, distally. We hypothesized that PWS would show reduced sensitivity to the distal rhythm and speech rate manipulations compared with matched controls.

Participants

Twenty-eight participants (22 males; 6 females), ranging from 18-52 years old, were selected for one of two groups: an experimental group of adults who stutter (n=15) and a control group (n = 13). All participants in both groups participated in Experiments 1 and 2 of this study. Participants were right-handed, monolingual native speakers of American English with normal hearing. Every participant was without concomitant developmental disorders (e.g., dyslexia, ADHD, learning delay, or other confirmed developmental or psychiatric conditions) and was not taking any medication affecting the central nervous system. Participants were matched on age, sex, years of education, working memory, receptive vocabulary (Peabody Picture Vocabulary Test, PPVT-4), expressive vocabulary (Expressive Vocabulary Test, EVT-2), and a measure of articulation (Goldman-Fristoe Test of Articulation, GFTA-2). Stuttering severity was assessed off-line by reviewing video recorded samples of speech elicited through monologue, dialogue, and reading sample with a certified Speech-Language Pathologist and/or trained Masters student. The Stuttering Severity Instrument (SSI-4) was used to assess frequency and duration of stuttering-like and other disfluencies occurring in the speech sample, as well as any physical concomitants associated with stuttering. These measures were then incorporated into a composite stuttering severity rating. Participants in the stuttering group ranged from very mild to severe in severity ratings from the SSI-4. Table 1 shows the means and standard deviations for age, handedness, education, years of musical training, expressive and receptive language measures, articulation measures, and working memory for both groups, and the stuttering severity for the stuttering group.

	Group		t-test
Measure	Control	Stuttering	p
Age	25 (7.42)	28.9 (11.36)	0.300
Handedness	80.77 (17.06)	78 (16.12)	0.663
Education	15.38 (1.39)	16.27 (2.25)	0.232
Years of musical training	5.15 (3.89)	3.8 (4.71)	0.419
Peabody Picture Vocabulary Test	107.15 (5.98)	109.4 (10.2)	0.493
Expressive Vocabulary Test	107.31 (9.6)	115.73 (11.32)	0.045
Goldman-Fristoe Test of Articulation	100.69 (1.44)	100.8 (1.66)	0.857
Operation Span Score (Working Memory)	40 (18.58)	38.8 (14.99)	0.851

Table 1: Participant Information. The table displays statistics of age, handedness, education, working memory, and speech and language measures. Independent samples *t*-tests revealed a significant difference between-subjects on the expressive language measure, in which the stuttering group displayed a group average standard score 8.42 points higher than the control group.

Experiment 1: Distal Intonation Rhythm

Methods

The first experiment was a replication and extension of Experiment 1A of Dilley and McAuley (2008). This study utilized distal prosodic manipulations of F0 and duration to examine the effect of distal prosodic cues in linguistic parsing. Eight-syllable target sequences, such as banker helpful [tɑɪ mɝ dɝ bi] were used; the final four syllables had ambiguous lexical structure and could be organized into words in more than one way (e.g., timer <u>derby</u>, tie murder <u>bee</u>). Items selected for this study corresponded to a subset of items from Dilley, Mattys and Vinke (2010), which were also used in Morrill et al. (2014). Experimental items utilized one of two different manipulation patterns of distal F0 that was expected to result in strong effects on the parsing of the final syllables into either a monosyllabic or disyllabic final word. Monosyllabic and disyllabic contexts were created by varying the fundamental frequency (FO) across the initial five syllables. The final, target portions of these sequences remained acoustically ambiguous (containing the same F0 pattern) for both monosyllabic and disyllabic contexts, resulting in the possibility of perceiving either a monosyllabic word (e.g., *bee*) or a disyllabic word (e.g., *derby*).

In one distal prosodic pattern, a repeated low (L)-high (H) alternation of F0 on distal syllables was predicted to lead to grouping of L and H pitch elements across the entire sequence, which was predicted to generate perception of a disyllabic word (e.g., *derby*); these are referred to as Disyllabic contexts. In the other distal prosodic pattern, a repeated high-low alternation of F0 was predicted to lead to grouping of H and L pitch elements across the entire sequence, which was predicted to generate perception of a monosyllabic final word (e.g., *bee*); these are referred to as Monosyllabic contexts. In more detail, for both the monosyllabic and disyllabic contexts, the final three syllables of each target sequence received a high (H), low (L), high (H) F0 pattern (one F0 target per syllable). The first five syllables of each target sequence varied in F0 pattern; the first five syllables of the disyllabic context items received a L1-H2-L3-H4-L5- pattern, with one F0 target, H or L, on each syllable and the monosyllabic pattern received a H1-L2-H3-L4- HL5- pattern with one F0 target for each of the first four syllables, and a fall in F0 from H to L on the fifth syllable. These distal prosodic context manipulations predicted the perceptual grouping of the final syllables in the target sequence, in which the Disyllabic context was predicted to cause a relatively stronger prosodic boundary to be heard before the penultimate syllable (e.g., der) than before the final syllable (e.g., bee), so that listeners would group the final two syllables as a single disyllabic word (e.g., derby). In contrast, the Monosyllabic context was predicted to yield the reverse, with a stronger prosodic boundary to be heard before the final syllable (e.g., bee) than before the penultimate syllable so that listeners would hear a final monosyllabic word (e.g., bee). The Disyllabic context yielded a (L1-H2-)(L3-H4-)(L5-H6-)(L7-H8) grouping of the final three syllable sequence, displaying a larger boundary before syllable 7 than syllable 8, thereby critically yielding a disyllabic final word report (e.g., derby). In contrast, the Monosyllabic context yielded a (H1-L2-)(H3-L4-)(HL5-) (H6-L7-)(H8) grouping of the final three syllables, displaying a larger boundary before syllable 8 than before syllable 7. -The pattern of responses by typical listeners was as predicted: there was a strong tendency for listeners to hear the same acoustic material as ending in disyllabic words in Disyllabic contexts but as ending in monosyllabic words in Monosyllabic contexts. . The expected response for the low-high intonation pattern of the experimental items is to perceive and report a disyllabic word (derby) as long as the listener is perceiving the externally cued distal intonation rhythm and internally carrying it throughout the final two target syllables. These results were expected to be seen within the control group, and seen as a higher difference between the reported number od disyllabic words and monosyllabic words for either disyllabic or monosyllabic contexts. We expected PWS to demonstrate a lesser difference between the proportions of reported disyllabic and monosyllabic final words.

Twenty experimental stimuli were provided aurally to each participant via headphones, with each auditory speech sequence containing two disyllabic, non-compound words continued with a string of four syllables [tɑɪ mɝ dɝ bi]. Both target and filler sentences were read as connected speech, using monotone F0. Resynthesized speech stimuli were then derived from these utterances using the pitch-synchronous overlap-and-add (PSOLA) algorithm (Moulines & Charpentier, 1990) as implemented in Praat software. Methods for recording are described in Dilley, Mattys and Vinke (2010).

Ten filler sentences were included in the experiment to serve as practice and disguise the lexical ambiguity present in target sentences. Fillers consisted of a mixture of monosyllabic and disyllabic words with unambiguous lexical structure in varying positions within the string; five of the filler sequences ended in a disyllabic word and the other five ended in a monosyllabic word. Each experimental item consisted of initial primary stress on the two disyllabic word sequence, e.g., channel dizzy, followed by a four-syllable sequence that could be grouped into words in more than one way.

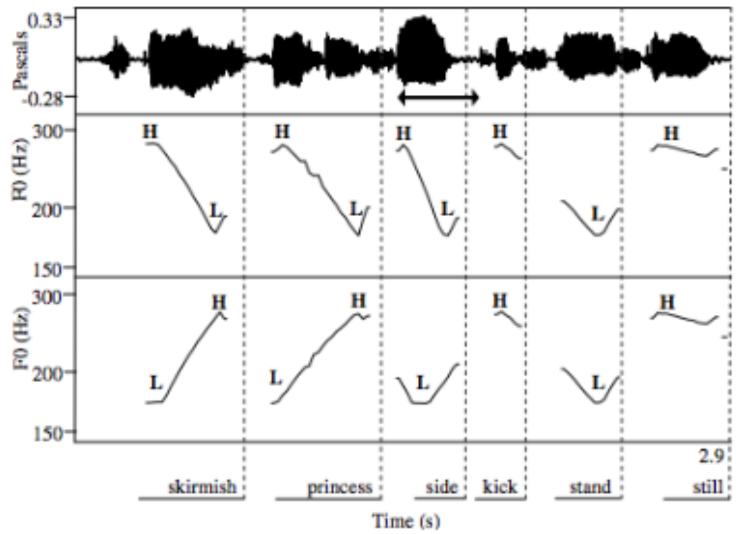

Figure 1: Distal Context Manipulations. Example of the distal context intonation pattern for stimuli used in this experiment. The top row displays a sample speech waveform of an auditory string. The bottom two rows show the fundamental frequency (F0) contour for the two different contexts (monosyllabic vs. disyllabic) along with the expected syllable parsing for each. H and L refer to high and low F0 targets, respectively. The middle row displays the intonation sequence and expected syllable parsing for the monosyllabic context, and the bottom for row for the disyllabic context. Note the targeted final two syllables for both contexts remain the same low-high pattern, and the lexical content remaining ambiguous with logical parsing for either a monosyllabic word (still) or a disyllabic word (standstill).

The distal prosodic context was expected to affect participants' proportions of disyllabic final words heard, although the portion of the string being judged consisted of identical acoustic material across conditions.

Participants provided a free recall typed report of the final word they heard for each sentence, with the dependent measure being the proportion of disyllabic responses reported for each experimental stimulus. Each participant parsed the final four syllables of the string into either a disyllabic final word (e.g., *timer* *derby*) or a monosyllabic final word (e.g., *tie murder* *bee*). All experimental items were presented in both a monosyllabic and disyllabic context, and were counterbalanced across participants. The entire experiment took about 12 minutes to complete.

Results

Measures were taken on the proportion of reported disyllabic final words as a function of type of word context (disyllabic vs. monosyllabic) and group (control vs. stuttering). The typed reported words for each participant were coded by two different raters as being either a monosyllabic or disyllabic word. There were no discrepancies between the two raters, and every participant's responses were included in the analysis.

Table 2 displays the means and standard deviations for both the disyllabic and monosyllabic words reported across the groups. A 2 (type of word context: monosyllabic vs. disyllabic) x 2 (group: control vs. stuttering) mixed-measures ANOVA on disyllabic response proportions was conducted. First, there was a significant main effect of Word Context ($F(1, 26) = 337.121$, $p < .001$, $\eta_p^2 = .928$), corresponding to a higher proportion of disyllabic responses in the disyllabic context ($M = .940$) than the monosyllabic context ($M = .164$). These results are consistent with the expected higher proportion of disyllabic responses due to effects of Word Context.

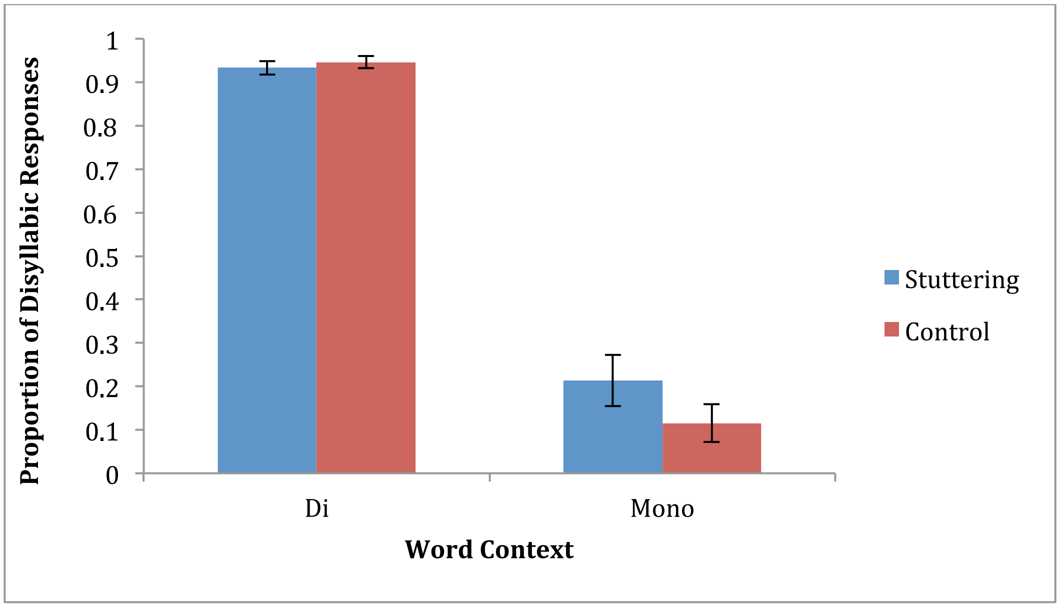

Figure 2: Proportion of Disyllabic Responses. The effects of Word Context within-subjects on proportion of reported disyllabic responses for both experimental stimuli contexts, is displayed above. Stuttering refers to the experimental group of adults who stutter, and control refers to group of typically-fluent speakers.

No significant effect of group ($F(1, 26) = 1.384$, $p = .250$, $\eta_p^2 = .051$) or interaction ($F(1, 26) = 1.720$, $p = .201$, $\eta_p^2 = .062$) was found.

Discussion

In Experiment 1, we examined how the stuttering group would respond to the manipulation of the distal prosodic cues in lexical segmentation of an ambiguous target sequence by measuring the proportion of reported disyllabic final words for each experimental stimulus. It was hypothesized that PWS would be less sensitive to this rhythmic cuing, resulting in a smaller difference in the proportion of disyllabic words in Disyllabic vs. Monosyllabic distal contexts compared with control participants. Our finding of no significant effect of group demonstrates that PWS showed similar sensitivity to rhythmic prosody compared with the control group. The significant main effect of Word Context within-subject for both groups indicates that both PWS and the controls exhibited sensitivity to distal prosodic cues, which influenced their syllabic parsing of the final target string. The design of this experimental paradigm relies on the sensitivity of the listener to not only perceive the distal prosodic cues of manipulated F0 and duration, but also internally generate the intonation rhythm and carry it through to the final four syllables in order to parse a disyllabic final word. The PWS group resulted in reporting a higher proportion of disyllabic words than monosyllabic words, similar to pattern exhibited by the control group. Thus, we can conclude that PWS have similar sensitivity to prosodic elements in their speech perceptions, as indexed by these distal prosody manipulations, as compared with fluent speakers. These results also suggest that PWS were able to internally generate the intonation rhythm perceived throughout the distal material in a manner that carried over to the lexical ambiguous targeted sequence; this suggests that under the present experimental conditions the internal rhythm generation of PWS was intact for a minimal time-span of carryover. Therefore, we did not find support for the hypothesis that PWS would demonstrate a lessened sensitivity to the distal prosodic cuing of intonation rhythm and duration.

Experiment 2: Distal Speech Rate

Methods

The second experiment used in this study was a replication and extension of the experimental paradigm used by Dilley and Pitt (2010). This study investigated sensitivity to prosodic cues at the syllable level by manipulating context speech rate. The experiment utilized 40 experimental items; each included a sentence that contained target syllables with a function word presented via auditory recordings. Each sentence had a function word that is optional, as the sentence will be grammatical with or without the function word. For example, in the sentence "Don must see the harbor (or) boats," the target region consisted of the function word *or*, the preceding syllable *–bor*, and the following phoneme [b]. The context region included the non-target portions of the context.

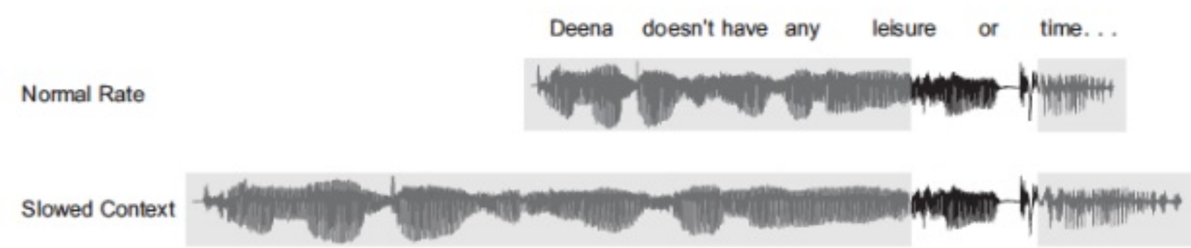

Figure 3: Distal Rate Manipulations. Waveforms of a sample altered stimulus exampling normal (unaltered rate) versus a slowed distal context (slowed rate). Sections of the waveform without background shadowing correspond to the target region, which consisted of a critical function word (e.g., "or") plus the preceding syllable and following phoneme, following Dilley and Pitt (2010).

The experimental stimuli used four different speech rate expansion factors (1, 1.3, 1.6, 1.9), each involving a temporal manipulation in which the distal context of the phrase is modified while the target syllables remain un-modified. For slowed rates, both the portion of speech preceding the target region and the region that followed it were multiplied by 1.3, 1.6, or 1.9 using the PSOLA algorithm in Praat (Dilley & Pitt, 2010). That is, the duration of the context was 130%, 160%, or 190% the duration of the original context, thus slowing the speech rate. The unaltered items were multiplied by 1.0 using the PSOLA algorithm, maintaining their original speech rate. It was expected that if the listener had sensitivity to the distal speech rate cues, he or she would report a successively smaller proportion of function words as the distal speech rate was made increasingly slow. The dependent measure was the proportion of reported function words. We predicted that if PWS exhibited a lessened sensitivity to distal speech rate compared with control participants, there would be a smaller drop in the proportion of function words with increasingly slow speech rates compared with controls. Each participant heard every experimental sentence and the pairing of speech rates with experimental items was counterbalanced across participants. Filler items had no ambiguity regarding the number of words they contained. They were presented at either the unaltered or slowed rate.

The participants were shown the first four words of the phrase on the computer screen and were asked to transcribe the remainder of the words they heard. If the participant showed sensitivity to the distal speech rate manipulations, he/she did not perceive the function word when the distal context was slowed, but did perceive the function word when the distal context was un-modified. Each participant's responses were coded by two different raters. The reported function word was determined by the following definition; a word (such as a preposition or a conjunction) that is used mainly to show grammatical relationships between other words. If the reported function word was not the same as the spoken function word in the auditory stimuli, it was still counted as a reported function word with the constraints that the immediate preceding and proceeding phonemic environments of the target context remained accurate.

Results

Analyses were based on whether participants reported hearing a function word in the acoustically ambiguous region for each utterance. The frequency of transcribing a function word in the target region was scored; presence of a critical function word in the target region was coded as 1, and absence of a function word was coded as 0. Responses that indicated inattention to the stimuli (i.e., did not include the preceding and/or following phonemic environments) were not included in the analysis (9%).

Figure 4 displays the proportion of reported function words as a function of rate and group. A 4 (Rate: 1, 1.3, 1.6, 1.9) x 2 (Group: control vs. stuttering) mixed-measures ANOVA on reported function word was conducted. There was a significant main effect of rate ($F(3, 78) = 118.230$, $p < .001$, $\eta_p^2 = .820$).

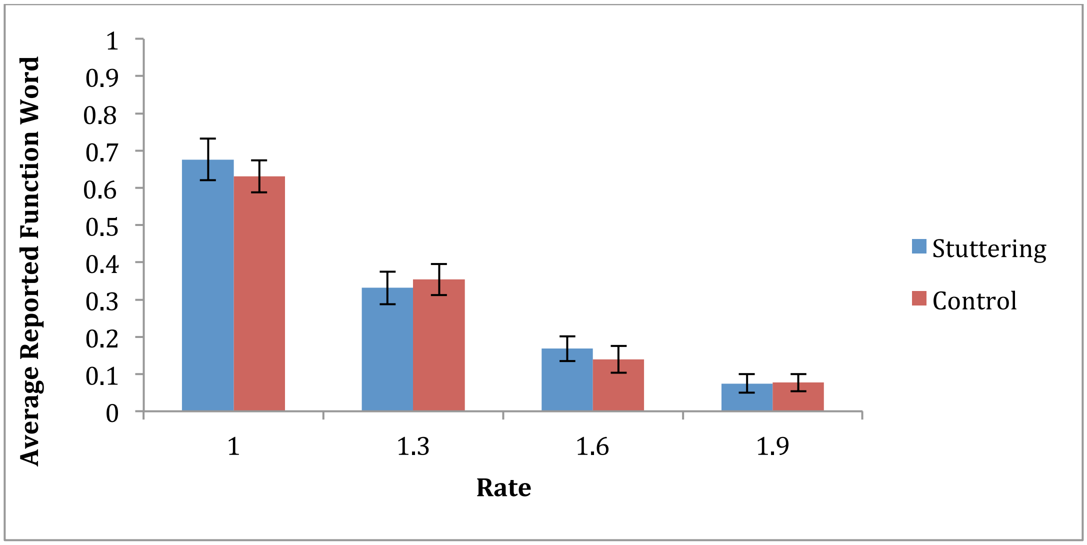

Figure 4: Reported Function Word by Rate. Average reported function word for each rate between-subjects displayed no significant differences between stuttering and control groups. Distal context rate had a significant main effect for both groups resulting in a gradual decline of average reported function words as the rate was slowed.

There was no significant main effect of group ($F(1, 26) = .032$, $p = .860$, $\eta_p^2 = .001$) or interaction ($F(1, 26) = 262.028$, $p = .820$, $\eta_p^2 = .019$) found. Given the significant effect of Rate, Bonferroni-corrected, post-hoc *t*-tests were conducted to further examine pairwise differences among the levels. Results demonstrated significant differences between all pairs of rate levels at $p < .001$ except for 1.6 vs. 1.9 ($p = .057$). These results demonstrate that the speech rate effect is robust across both PWS and control participants.

A few additional observations are noteworthy about this data. First, in the normal-rate condition, function word reports were quite high; the fact that they were not at ceiling is expected given that the speech was casually spoken. Moreover, note that a comparison of reports in the normal-rate condition and the slowed-context conditions (1.3, 1.6, and 1.9) showed that merely slowing the context surrounding a function word caused the rate of reported function words to drop. The degree of drop was shown to be consistent across both groups. These results display a gradual decrease in average reported function word as a function of slowed distal rate outside of the final two rates, in which the decrease of the reported function word means was not significant.

Discussion

Experiment 2 showed a significant main effect of distal speech rate, but PWS and control participants showed similar sensitivity to distal speech rate, resulting in no significant effect of group. These results did not support our original hypothesis that PWS would exhibit a lessened sensitivity in perception of distal speech rate, and therefore would not have as large of a difference of reported frequency words between distal speech rates as the control population would. Making the duration of a stretch of speech containing a function word fast relative to its slowed distal context affected the number of morphophonological units perceived by the participants in both groups to a similar degree. We obtained clear evidence that an alteration of speech rate on the distal context affected the word segmentation of the target context, extending prior work by Dilley and Pitt (2010) to a clinical group of PWS. Due to the lack of group differences in the comparison of PWS and controls, we can infer that PWS exhibited sensitivity to distal speech rate and did not demonstrate impairment in their perception or internal generation of the distal rate throughout the stimulus.

General Discussion and Conclusions

Stuttering is characterized as a neurodevelopmental disorder that impacts the rhythmic flow and timing of spoken language. Despite the amount of previous research and plausible theories on the etiology of stuttering, there still is no generally accepted cause of the underlying mechanisms affecting fluency. Prior evidence has shown that fluency-enhancing techniques for PWS utilize some type of externally-generated rhythm or timing including singing, choral reading, or using a metronome, suggesting that internal beat and rhythm generation is problematic for PWS (Adams & Ramig, 1980; Ingham & Carroll, 1977; Glover, Kalinowski, Rastatter, & Stuart, 1996; Wingate & Howell, 2002). Other research supported this idea with evidence that children who stutter displayed lessened white matter activation in the basal ganglia thalamocortical (BGTC) network, which involves the neural cortices that are utilized in the auditory perceptions of rhythm, beat, and timing (Chang & Zhu, 2013; Grahn & McAuley, 2009;Toyomura et al., 2015). While stuttering is typically thought of a speech production disorder, studies have supported that speech productions of people who stutter (PWS) can be influenced by the perception of prosodic aspects of their own speech (Hargrave, Kalinowski, Stuart, Armson, & Jones, 1994; Wingate, 1966). This study was thus conducted to better understand the nature of the possible deficits in the perceptions of rhythm and timing and internal beat generation in persons who stutter. We investigated the perception of distal speech rhythm and rate in a group of adults who stutter and a control group.

In Experiment 1, distal intonation rhythm was manipulated and both groups responded with the final word that they heard. The experimental manipulations utilized distal pitch pattern which were expected to cause differential perception of disyllabic vs. monosyllabic final words if they were sensitive to distal cues and could use them to generate a rhythm (Dilley & McAuley, 2008). Our hypothesis proposed that the stuttering group would show less sensitivity to these prosodic manipulations, which was predicted to manifest as a smaller difference in disyllabic final word reports in disyllabic vs. monosyllabic context conditions. In Experiment 1, PWS reported a higher proportion of disyllabic responses in disyllabic contexts and similar effects of context type as the control group, indicating that they perceived the distal prosodic manipulations and internally generated a rhythm that influenced the syllable parsing of the final word in a manner that was similar to controls.

In Experiment 2, distal rate of speech was manipulated and both groups reported the final words, with the dependent variable being function word reports. The experimental manipulations utilized a slowing of distal speech rate to cause the function word to be heard increasingly less often with increasingly slow distal speech rates. Our hypothesis proposed that PWS would show a lessened sensitivity to the distal prosodic cues, resulting in a smaller decrease in function word reports as distal speech rate was slowed, compared with controls. In Experiment 2, the significant main effect of rate in PWS indicated that PWS were sensitive to the distal manipulations of rate, since they reported fewer function words as distal was slowed, mirroring the pattern for controls.

Based on our hypothesis, if PWS have a deficit with internally generating rhythm and timing, we might expect that they would have difficulty generating an internal structure from perceived speech rhythmic cues that would manifest as a difference between control and PWS in a task of distal prosody. Both Experiments 1 and 2 showed that PWS performed similarly in these tasks to the control group. In Experiment 1, the significant main effect of word context for both groups with no significant effect of group stood in contrast to the hypothesis of lessened sensitivity to the distal prosodic manipulations of intonation rhythm in PWS.

These results are inconsistent with the idea that the ability to generate prosody perceptually is not intact in PWS. That is, PWS did not demonstrate an impaired function within their perceptions of rhythm and timing cues in distal speech or in their ability to internally generate rhythms and rates cued externally throughout the ambiguous and unaltered targeted speech strings. This conclusion on the surface appears inconsistent with proposals of deficits in rhythm perception caused from dysfunction within the basal ganglia and the BGTC network (Grahn, 2009). Evidence of a deficit in rhythm perception was recently found by Wieland (2015) in children who stutter (CWS). This evidence is also supported by the reduced white matter tracts found in CWS by Chang (2013) compared with controls; these tracts are expected to affect connectivity in the BGTC network.

The failure to find evidence for deficits in perception of rhythmic aspects of speech and apparent conflict with findings from previous literature could be explained with reference to four proposals. The first proposal is that the deficits in rhythm perception found by Wieland (2015) were due to the difference in developmental stage of the populations studied: children who stutter (CWS) for Wieland (2015), in contrast to adults who stutter for the present study. Under this explanation, deficits in rhythm perception exhibited in CWS may disappear by adulthood, possibly due to the development of a compensatory strategy. A second possibility for the contrasting results of the present study is due to the robust strength of the manipulations utilized in this study. There is a possibility that deficits in rhythm and timing perception in PWS might be subtle; thus, such deficits might not have been observed in the present study due to the large effect sizes produced by the manipulations across both paradigms used in this study. The third possible reason for the contradictory results is that this study utilized perception of rhythm and timing in speech, while Wieland et al. (2015) used a perception task based solely on tones, without speech. It is possible that the difference in tasks (tones vs. speech) could have resulted in PWSs' relying on a different set of characteristics for speech vs. tones. The fourth proposal for the contradictory results across studies could be that the manipulations utilized in this current study did not rely heavily enough on the internal generation of rhythm and timing. The generation of the rhythms and timings utilized in these experiments were cued externally through the prosodic manipulations of distal contexts and carried throughout a minimal time-span, rather than fully being generated internally. Thus, the present results are consistent with the hypothesis that PWS can perceive rhythms and timings cued by externally provided distal prosodic cues but that they have deficits in internal rhythm and timing. Future research should address this and further investigate with a paradigm utilizing purely internally-generated rhythm and timing information.

It has been shown that internal generation of rhythm and pacing in people who stutter (PWS) can be influenced by perception of prosodic aspects of their own speech (Hargrave, Kalinowski, Stuart, Armson, & Jones, 1994; Wingate, 1966). Guenther (1994) suggested that auditory-motor integration is used for spoken language of fluent speakers and that speech production occurs by integrating perceptions of self-speech into ongoing speech production. Previous research points to a problem in some aspect of the auditory-integration process in a manner that generates disfluency. There is evidence that shows PWS have difficulty with this auditory-motor integration of speech perceptions into speech production (Cai et al., 2014). These findings of this study help us to distinguish among three possible loci for the deficit in PWS related to the auditory-motor integration of prosody in speech identified by Cai et al. (2014) and related research. One is the possibility of difficulty in the perception of prosody. A second possibility is difficulty in the production of prosodic elements in speech. A third possibility is difficulty of integrating perception of prosodic elements from self-speech into speech productions. These possibilities are discussed below with reference to the findings of the present study.

While stuttering is thought of as being a problem in the production of speech, speech perception has been shown to play an important role in the speech production of PWS (Hargrave, Kalinowski, Stuart, Armson, & Jones, 1994; Wingate, 1966). This is evidenced by the effects of changing the auditory feedback of pitch or timing on the fluency of PWS with delayed auditory feedback (DAF) devices (Cai et al., 2014). The possibility of a problem with perception of prosodic elements of speech provides one possible explanation for the underlying deficits in PWS. The results of the present study suggest that the perception of rhythm and timing in adults who stutter remains intact. These findings suggest that the problem does not seem to lie within the perceptions of prosodic elements of speech in PWS.

The hypothesis of difficulty in speech production is another potential locus of influence of prosody in underlying deficits of PWS. If the problem is within the motor mechanisms for speech production, PWS would not be able to produce speech that exhibited prosodic elements such as appropriate rhythm, timing, or pitch. PWS can however demonstrate appropriate prosodic elements in their speech productions with the assistance of externally generated rhythm and pacing devices (Hargrave, Kalinowski, Stuart, Armson, & Jones, 1994; Wingate, 1966). The fact that PWS can exhibit prosodic elements in fluent speech with assistive fluency-enhancing devices leads to the idea that the problem does not rely in solely the production of these prosodic elements. Further evidence was presented in a recent poster presentation at the annual conference of the American Speech-Language-Hearing Association by Wieland et al. (2014) that showed that prosodic aspects of fluent speech in children who stutter were similar to those of a group of matched control children. These results suggest that an underlying deficit in production of prosodic elements such as rhythm and timing are not likely to be a core deficit in PWS.

The third possibility is that the problem lies within the integration of perception into speech production. This possibility is supported by the results found by Cai et al. (2014) showing a deficit in PWS in the integration of perturbed auditory feedback into their speech productions. Cai et al. (2014) demonstrated results of weak error correction of speech production in PWS when their perceptions of timing had been disrupted. This idea of difficulty within auditory-integration is supported by the fluency-inducing effects of auditory-feedback devices that alter the pitch and timing in auditory feedback of self-speech. Without altered auditory-feedback, the incorporation of normal feedback into speech production seems to be disrupted, as it typically results in disfluency. Once prosodic elements of the auditory feedback is manipulated, fluency is often enhanced, supporting the idea of a disruption in the integration of auditory-feedback (perception of self-speech) into speech production.

From previous literature and the results of this study, the perception and production of prosodic elements of speech, each considered individually, do not display a problem underlying the deficits of PWS. Thus, the present work supports the idea that a new, fruitful hypothesis to be tested involves investigating how prosodic information from speech perception is integrated within speech production. While this study primarily focused on perception of prosodic elements and internal rhythm generation in PWS, a focus of research on the auditory-integration of speech perception into motor production of speech is suggested to be most productive for investigation of underlying deficits in PWS associated with prosody. Despite the lack of understanding of overall core deficits in PWS, the results from the current study have provided evidence of an intact ability of PWS to perceive prosodic elements in speech and to internally generate rhythm and timing following distal prosodic cues in speech. Research defining underlying deficits (or the lack thereof) have significant clinical implications, since core strategies or techniques can be recognized when core deficits or strengths within a clinical population are recognized and understood. . The results of this study shed light onto new potential areas for research to continue investigation of an underlying deficit related to prosody in PWS.

APPENDICES

Appendix A: Experiment 1 Items List

1. banker helpful (tie murder bee/timer derby)
2. kettle heaven (Tim burrow bow/timber oboe)
3. pebble dollar (bar lever chew/barley virtue)
4. gossip oyster (pan treaty coy/pantry decoy)
5. angry index (lay birdie fence/labor defense)
6. chapter elbow (rue beaver gin/ruby virgin)
7. kitchen dealer (may beanie grow/maybe negro)
8. hero vacuum (sell early gull/cellar legal)
9. bullet junior (come feeding key/comfy dinky)
10. liquid perish (broad leasing king/broadly sinking)
11. lumpy danger (chair eager knee/cherry gurney)
12. plasma honey (pigs typo low/pigsty polo)
13. blanket mounted (ham mercy nick/hammer scenic)
14. tourist robin (draw musty plea/drama steeply)
15. trouble wealthy (limb burner sing/limber nursing)
16. nature lazy (faux meaty tour/foamy detour)
17. lady jacket (bran diesel tree/brandy sultry)
18. husband lemon (fan seaman cheese/fancy munchies)
19. center northern (two cancer plus/toucan surplus)
20. mixture pleasure (class seedy pose/classy depots)

Appendix B: Experiment 2 Items List

1. The murrays are (our) favorite
2. The johnsons are (our) rich relatives from California
3. The accountants are (our) wise advisors
4. We won't have any winter (or) wet weather
5. Marty gave him a dollar (or) twenty last week
6. Fred would rather have (a) summer or lake
7. Mom said these are (our) gray gloves
8. Sam got it in writing before (her) last wishes
9. It's not long before (her) rare wit
10. John didn't tell the junior (or) representative about it
11. The snowsuits are (our) only
12. The Clines are (our) good friends
13. People were offended after (her) rude
14. These copy machines are (our) largest
15. Anne wanted to see (a) very funny
16. Susan said those are (our) black socks
17. New laws are (our) last chances for a change in climate
18. Deena doesn't have any leisure (or) time
19. Connor knew that bread and butter (are) both
20. It's not easy to convey (a) likely
21. Sam knew there (are) apples
22. Todd said there (are) rooms
23. John said he would obey (a) rebel
24. Sally might try (a) liquid
25. The sign was replaced after (her) black
26. Don must see the harbor (or) boats
27. The message was clear after (her) blank
28. Tina decorated with paper (or)lace in most rooms
29. Sue said there (are) lunches
30. Phil and Mary are (our) young cousins
31. Glenn thought his friend and neighbor (are) like plenty
32. Dave asked how long it takes to repay (a) large
33. Rose knew that there (are) lamps
34. Dawn said that it's easy to go to (a) regular
35. Those tickets are (our) late entries
36. John called her sugar (or) honey most of the time
37. The leaves fell after (her) green
38. They were sad after (her) poor
39. The Lees are (our) old friends from high school
40. It's not long before (her) bad

REFERENCES

REFERENCES

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: a critical review of possible relations. *Journal of communication disorders*, *37*(4), 325-369.

Adams, M. R., & Ramig, P. (1980). Vocal characteristics of normal speakers and stutterers during choral reading. *Journal of Speech, Language and Hearing Research, 23*(2), 457-469.

Bergmann, G. (1986). Studies in stuttering as a prosodic disturbance. *Journal of Speech, Language and Hearing Research, 29*(3), 290-300.

Besozzi, T. E., & Adams, M. R. (1969). The influence of prosody on stuttering adaptation. *Journal of Speech, Language, and Hearing Research, 12*(4), 818-824.

Cai, S., Beal, D. S., Ghosh, S. S., Guenther, F. H., & Perkell, J. S. (2014). Impaired timing adjustments in response to time-varying auditory perturbation during connected speech production in persons who stutter. *Brain and language*, *129*, 24-29.Chang, S.-E., & Zhu, D. (2013). Neural network connectivity differences in children who stutter. *Brain, 136*(12), 3709-3726.

Coull, J.T., 2004. fMRI studies of temporal attention: allocating attention within, or towards, time. Cogn. Brain Res. 21, 216–226.

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

Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language, 59*(3), 294-311.

Dilley, L. C., & Pitt, M. A. (2010). Altering context speech rate can cause words to appear or disappear. *Psychological Science, 21*(11), 1664-1670.

Glover, H., Kalinowski, J., Rastatter, M., & Stuart, A. (1996). Effect of instruction to sing on stuttering frequency at normal and fast rates. *Perceptual and Motor Skills, 83*, 511-522.

Grahn, J.A., & Brett, M., 2007. Rhythm perception in motor areas of the brain. J. Cogn. Neurosci. 19, 893–906.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *Neuroimage, 47*(4), 1894-1903. doi:10.1016/j.neuroimage.2009.04.039

Guenther, F. H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics, 72*(1), 43-53.

Hall, K. D., Amir, O., & Yairi, E. (1999). A longitudinal investigation of speaking rate in preschool children who stutter. *Journal of Speech, Language and Hearing Research, 42*(6), 1367–1377.

Hargrave, S., Kalinowski, J., Stuart, A., Armson, J., & Jones, K. (1994). Effect of frequency-altered feedback on stuttering frequency at normal and fast speech rates. *Journal of Speech, Language, and Hearing Research, 37*(6), 1313-1319.

Harrington, D.L., Haaland, K.Y., 1999. Neural underpinnings of temporal processing: a review of focal lesion, pharmacological, and functional imaging research. Rev.Neurosci. 10, 91–116.

Harrington, D.L., Haaland, K.Y., Knight, R.T., 1998b. Cortical networks underlying mechanisms of time perception. J. Neurosci. 18, 1085–1095.

Hazeltine, E., Helmuth, L.L., Ivry, R., 1997. Neural mechanisms of timing. Trends Cogn. Sci. 1, 163–169.

Ingham, R. J., & Carroll, P. J. (1977). Listener judgment of differences in stutterers' nonstuttered speech during chorus-and nonchorus-reading conditions. *Journal of Speech, Language and Hearing Research, 20*(2), 293-302.

Kent, R. D. (1984). Stuttering as a temporal programming disorder. In R. F. Curlee & W. H. Perkins (Eds.), *Nature and treatment of stuttering: New directions* (pp. 283-301). San Diego: College-Hill Press.

Morrill, T. H., Dilley, L. C., & McAuley, J. D. (2014). Prosodic patterning in distal speech context: Effects of list intonation and f0 downtrend on perception of proximal prosodic structure. *Journal of Phonetics*, *46*, 68-85.

Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. Speech Communication, 9(5-6), 453–467.

Nenadic, I., Gaser, C., Volz, H.-P., Rammsayer, T., Hager, F., Sauer, H., 2003. Processing of temporal information and the basal ganglia: new evidence from fMRI. Exp. Brain Res. 148, 238–246.

Packman, A., Code, C., & Onslow, M. (2007). On the cause of stuttering: Integrating theory with brain and behavioral research. *Journal of Neurolinguistics, 20*(5), 353-362. doi: 10.1016/j.jneuroling.2006.11.001

Packman, A., Onslow, M., Richard, F., & Van Doorn, J. (1996). Syllabic stress and variability: A model of stuttering. *Clinical Linguistics & Phonetics, 10*(3), 235-263.

Rao, S.M., Mayer, A.R., Harrington, D.L., 2001. The evolution of brain activation during temporal processing. Nat. Neurosci. 4, 317–323.

Salverda, A. P., Dahan, D., Tanenhaus, M. K., Crosswhite, K., Masharov, M., & McDonough, J. (2007). Effects of prosodically-modulated sub-phonetic variations on lexical competition. Cognition, 105,466-476.

Schubotz, R.I., Friederici, A.D., von Cramon, D.Y., 2000. Time perception andmotor timing: a common cortical and subcortical basis revealed by fMRI. Neuroimage 11, 1–12.

Toyomura, A., et al., Effect of an 8-week practice of externally triggered speech on basal ganglia activity of stuttering and fluent speakers, NeuroImage (2015), http://dx.doi.org/10.1016/j.neuroimage.2015.01.024

Wieland, E. A., Dilley, L. C., Burnham, E. B., Redford, M., & Chang, S-E. (2014). Prosodic characteristics of fluent speech by children who do and do not stutter. Poster presented at the Annual Meeting of the American Speech-Language Hearing Association, Orlando, FL, November 20-22, 2014

Wieland, E. A., McAuley, J. D., Dilley, L. C., & Chang, S. E. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain and language*, *144*, 26-34.Wingate, M. E. (1966). Prosody in stuttering adaptation. *Journal of Speech, Language, and Hearing Research, 9*(4), 550-556.

Wingate, M. E., & Howell, P. (2002). Foundations of stuttering. *J Acoust Soc Am*(112), 1229-1241.

World-Health-Organization. (2010). International statistical classification of diseases and related health problems. Retrieved 4/30/13, from http://apps.who.int/classifications/icd10/browse/2010/en#/F98.5