FUNDAMENTAL FREQUENCY CHARACTERISTICS AFFECT JUDGMENTS OF SEGMENTAL, SUPRASEGMENTAL, AND INDEXICAL PROPERTIES OF CHILDREN WHO STUTTER

By

Danielle Gage

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

FUNDAMENTAL FREQUENCY CHARACTERISTICS AFFECT JUDGMENTS OF SEGMENTAL, SUPRASEGMENTAL, AND INDEXICAL PROPERTIES OF CHILDREN WHO STUTTER

By

Danielle Gage

Stuttering is a speech and language disorder characterized by disruptions in the production of speech sounds. Stuttered speech often includes repetitions of words or parts of words, prolongations of speech sounds, blocks or pauses, interjections such as “um”, or “like”, or delays of initiating sounds. A critical question from a theoretical standpoint is whether the perception of prosody within a speech signal is independent of the perception of other components of the speech signal or interdependent with other components. Prior findings have indicated that the speech of typically developing child talkers, when modified to a lower fundamental frequency (F0) level, was judged to be slower, less fluent, and less intelligible than speech with a higher F0. The present study tested the generalizability of these findings to speech of children who stutter. F0 of speech of children who stutter was raised, lowered, or kept at its original level. Listeners assessed anxiety, cognitive abilities, fluency, gender, likeability, speech rate, speech and language abilities based on an 8 point scale. In addition, listeners estimated the age of the speaker, and the percentage of intelligible words. Results are discussed in terms of theories of speech perception and the practical implications for stuttering therapy.

ACKNOWLEDGMENTS

I would like to thank my thesis committee member, Dr. Laura Dilley, Ph.D, Dr. Soo-Eun Chang, Ph.D, CCC-SLP, and Dr. Eric Hunter, Ph.D, for all of their support with this research. Additionally, I thank Elizabeth Wieland, Amanda Depalsmaker, Kayla Tillman, Stephanie Schmidt, and other members of the MSU Speech Perception-Production Laboratory and MSU Speech-Neurophysiology Laboratory for their assistance with this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

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

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

INTRODUCTION

Stuttering is a speech disorder characterized by disruptions in the production of fluent speech, otherwise known as disfluencies (ASHA, 2014). These disfluencies can include repetitions of words or parts of words, prolongations of audible or inaudible (i.e., block) speech sounds, interjections such as "um" or "like", or unintentional delays of initiating sound (ASHA, 2014). Stuttering is often described as a timing disorder, disintegration or discoordination of subglottic and glottal processes, and of articulation with phonation and respiration (Adams, 1975; Bergmann, 1986; Perkins, 1981; Perkins, Rudas, Johnson, & Bell, 1976; Van Riper, 1982). These disruptions may arise from a number of possible mechanisms related to aberrant speech motor control and/or its interactions with linguistic, cognitive and emotional processes (Kleinow & Smith, 2000; Smith & Kelly, 1997; Tasko, McClean, & Runyan, 2007). Stuttering affects approximately 5% of children and 1% of the general population; however, the etiology of stuttering is yet unresolved.

Stuttering plays a significant role in the perceived effectiveness of an individual's communication. There is reasonable evidence that indicates that children who stutter are likely to have concomitant language, articulation, and phonological concerns (Bloodstein, 1995; St. Louis & Hinzman, 1988; Wolk, Conture, & Edwards, 1990). In fact, 30-40% of children who stutter also exhibit phonological and/or articulatory disorders (Wolk, Conture, & Edwards, 1990; Yaruss & Conture, 1996). Anderson, Wagovich, and Hall (2006) found that children who stutter display an asynchrony in speech-language processing systems (i.e., speech-language dissociation), which may imply slower phonological processing, resulting in increased phoneme errors. Increased phonemic errors, as seen in PWS, may be perceived as a breakdown in communication. A study by Von Tiling (2011) found that PWS are afraid of being viewed as "communicatively incompetent" due to perceived difficulties in making precise points and arrangements in speech, as well as avoiding and correcting for misunderstandings. The ability to adapt messages effectively and appropriately to communicative interaction contexts is generally perceived to be hindered in PWS, which may further indicate a deficit in overall effective communication (Spitzberg & Cupach, 1984; Rickheit, Strohner, & Vorwerg, 2008).

Moreover, stuttering often leads to negative psychosocial consequences such as lack of confidence, self-assertion and general self-advocacy skills, as well as lack of vitality, social, emotional, and mental functioning, which may commonly result in elevated levels of anxiety (Bricker-Katz, Lincoln, McCabe, & 2010; Craig, Blumgart, Tran, 2009; Craig & Tran, 2006; Craig, Tran, Wikesuriya, & Boord, 2006). Individuals with any kind of overt disorder, such as stuttering, are often perceived with more negative connotations. In addition to negative perceptions, PWS also face negative sanctions as well (i.e pity, condescension, exclusion) (Klassen, 2001). Research by Gabel, Blood, Tellis, and Althouse (2004) reported that stuttering negatively affected the perceptions of appropriate career choices for people who stutter, demonstrating that people who stutter (PWS) were advised against becoming an attorney, a judge, a minister, or a psychologist. These researchers also found that PWS often remain constrained by the impact that stuttering has on his/her communication, and that they frequently struggle with a fear of speaking, as well as a fear of negative evaluation by others. Therefore, it is common for PWS to apply learned and compensatory techniques to assist in fluent speech production, which can then decrease fear of speaking situations and increase overall confidence (Bricker-Katz, Lincoln, & McCabe, 2010).

Stuttering can be assessed in a variety of ways either quantitatively or qualitatively (i.e., impressionistically or perceptually). Quantitative assessment of stuttering can be accomplished by trained Speech-Language Pathologists using the Stuttering Severity Instrument (SSI), by counting the frequency of disfluencies per hundred words or syllables, and the duration of disfluencies. Quantitative scales have been shown to produce highly reliable results, and often indicate that fluent speakers stereotype PWS in similarly consistent ways (Hughes, Gabel, Irani, & Schlagheck, 2010). However, qualitative judgments are necessary to some degree even when administering quantitative assessments, as clinical judgments are incorporated. Therefore, qualitative judgements may often provide for a more functional and holistic assessment, as such depends jointly upon overlapping speech components, primarily in children who stutter (Ferrer, L., Scheffer, & Shriberg, 2010). Often, the foundation of many stuttering treatment approaches include techniques whose replicability and implementation depends largely upon clinician insight and judgment (Ingham, 1975; Ingham & Lewis, 1978; Ingham, Montgomery, & Ulliana, 1983). Purely qualitative studies have been rare, but may have more validity and are generally able to provide more information regarding listener perceptions and attitudes toward PWS (Hughes, Gabel, Irani, & Schlagheck, 2010).

The effective use of overlapping speech cues are critical to understanding the intentions of speech throughout communicative interactions, particularly when working with clinical populations such as children who stutter (Beilby & Byrnes, 2012). Much research has noted a relationship between the segmental, suprasegmental, and indexical components of speech, as they all simultaneously interact during speech. First, segmental aspects of speech refer to the organization of sequences of phonetic or syllabic components, relying on the lexical elements of the speech signal, otherwise referred to the gross communicative message (Kent, 1983). Second, suprasegmental, or prosodic, aspects of speech, are commonly characterized by the rhythm and intonation patterns of speech which are used to modulate and enhance the meaning of a spoken speech signal, direct attention to, and identify personal aspects of the speaker (Grossman, Bemis, Skwerer, & Tager-Flusberg, 2010; Darwin, 1975; Pallier, 1996). Third, indexical aspects of a speech signal refer to properties of an individual's speech variability and sociolinguistic identity (e.g., gender, age, race, class, speaking style), and group membership (e.g., normal talker, or talker with a speech-language disorder). Indexical aspects of speech can also refer to differences in vocal tracts (Peterson & Barney, 1952), glottal waves (Monsen & Engebreston, 1977), articulatory dynamics (Ladefoged, 1980), and native dialects (cited in Goldinger, 1998). In combination, these three aspects of speech (segmental, suprasegmental, indexical) are the constitutive components of speech signals. Therefore, it is of interest as to how the properties of segmental, suprasegmental, and indexical attributes overlap in order to consequently impact the perceptions of listeners in everyday situations.

The perception of prosody within a speech signal has been viewed by many researchers as independent of other components of the speech signal. Many theories of speech processing implicitly or explicitly claim independence of segmental, suprasegmental, and indexical properties, thus implying that it is possible to extract one aspect in order to accurately assess overall communicative competence. In fact, Halle (1985) and Joos (1948) refer to this process as "speaker normalization", which proposes that all suprasegmental and indexical aspects are filtered out of the segmental stream. Remez, Fellowes, and Rubin (1997) characterized this position by saying that in order to identify a word, the listener discards acoustic aspects in order to form an abstract representation of the linguistic message. Likewise, to identify speaker-specific information, the listener discards acoustic information specific to phonemes to form a representation of voice quality. Joos (1948) further suggested that voice details can be considered "noise", and are therefore irrelevant to phonetic perception. Therefore, suprasegmental and indexical aspects of the signal such as voice quality, utterance speed, and other circumstantial information are irrelevant when determining lexical and semantic content, and that it is rare for listeners to truly recall the salient acoustic properties of a speech signal (Halle, 1985). This idea was further supported by Studdert-Kennedy (1976) who proposed that since spoken word perception is seemingly unaffected by subtle variations in the speech signal, lexical entries are more abstract, and therefore unrelated. Pisoni (1993) refuted this idea suggesting that abstract representations enable lexical interpretation, providing robust, context-insensitive perception. Such theories assume that surface information, such as suprasegmental and indexical, aspects are filtered in speech perception.

By contrast, a growing body of literature supports the notion of interdependencies among suprasegmental, indexical, and segmental components in the perception of a speech signal (Pierrehumbert, 2003, Psioni, 1993). Such research states that perceived difficulties in segmental aspects of communication, such as those suggested for PWS, can extend into the effective use of prosodic characteristics and interact with indexical components of speech. Recent research has speculated that PWS have "unstable speech systems" that may be interrupted by variability within the speech signal. Packman, Onslow, Richard, and Doorn (1996) explored whether altering the prosodic aspect of syllable stress could reduce segmental stuttering behaviors. It was found that when PWS decreased the variability within the speech signal, stuttering behaviors decreased as well. This notion supports the relationship between speech motor execution (segmental), speech variability (indexical) and syllabic stress (suprasegmental) in providing a comprehensive speech signal for PWS. Bergmann (1986) suggested that suprasegmental effects of speech may be closely related to timing (segmental) within a speech signal, especially in PWS. This was supported by Riley-Graham (2011) who demonstrated that PWS have increased difficulty producing the intended phonemes with the proper suprasegmental properties of timing, rhythm, and pitch. Furthermore, Ladd, Silverman, Tolkmitt, Bergmann, and Scherer (1985) supported this idea of overlapping components stating that emotional and attitude dependent (indexical) factors affect overall intonation patterns and F0 values. Furthermore, changes in F0 could be potentially anchored by segmental components and structure within the overall speech signal, suggesting that segmental and suprasegmental information align when conveying speech signals (Arvaniti, Ladd, & Mennen, 1998; Ladd, Faulkner, Faulkner, & Schepman,1999, Bruce, 1977). In general, such studies have indicated a dependence between suprasegmental components and speech rate. Bergmann (1986) further speculated regarding the relationship of indexical components stating that 'person-dependent' (indexical) and 'communication-related' (i.e., communication demands, expectations) psychological factors also significantly impact motor execution. An additional study by Sommers and Barcroft (2006), hypothesized that differences among indexical components of speech (i.e, speaking style), would interact with lexical components of speech. Considered with previous findings, the results largely indicated that variability in indexical, suprasegmental, and segmental components of speech do significantly impact identification of spoken words (Mullennix, Pisoni, & Martin, 1989; Sommers, Nygaard, & Pisoni, 1994).

In general, aspects of speaker variation (indexical) interact with presentation rate (segmental), giving rise to the potential of affecting listener's perceptual identification of the speech context and of the speaker themselves (Goldinger, Pisoni, & Logan, 1991; Klatt & Klatt (1990). Because of such difficulties, many stuttering therapies focus on desensitizing PWS to reactions and attitudes of listeners, therefore suggesting that listener perceptions are pertinent for enhancing fluency (Johnson, 1959; Woods, 1978; Van Riper, 1982). A study by Ingham, Martin, and Haroldson (1985) investigated whether speech naturalness could be used to not only assess speech quality, but to modify speech quality. These researchers stated that speech naturalness scores based on perceptual ratings from naïve listeners could be used as functional contingencies, and could potentially lead to new approaches to stuttering therapy. Therapy approaches also often attempt to produce stutter-free speech by introducing new fluency-enhancing or modification patterns of speech. Such strategies include, but are not limited to whispering, choral speech and singing, paced/metronomic speech (Wingate, 1969, 1970; Stager, Jeffries, & Braun, 2004; Ingham, Bothe, Wang, Purkhiser, & New, 2012). However, therapy approaches consequently introduce speech patterns that are audibly different from the speech of individuals who do not stutter (Franken, Boves, Peters, & Webster, 1992). These therapy approaches correlate with Wingate's (1969, 1970) previously proposed a "vocalization hypothesis", which states that various forms of novel speech (e.g., whispering, singing, paced speech), create enhanced fluency for PWS due to alterations in the manner of vocalization. Wingate further hypothesized that fundamental frequency is one such alteration of vocalization, and such a change could lead to increased fluency patterns in PWS. A study by Healey (1982) sought to justify Wingate's hypothesis and identify certain parameters of speaking fundamental frequency (F0) associated with people who do and do not stutter. Results indicated that the F0 for PWS was slightly lower and yielded more standard deviations than normally fluent speakers. The F0 range, rate, and mean were significantly greater in PWS than in non-stuttering groups (i.e. suprasegmental aspects). The reduction of F0 frequency variability across studies is one speech aspect that can be objectively measured when comparing speech-production differences between PWS and fluent speakers.

Since specific segmental properties of a speech signal (e.g., articulation rate) are often targeted in speech therapy for stuttering and other disorders, it is of interest to determine the extent to which one informational component of the signal (e.g., the segmental lexical content) might be influenced by other informational components of the acoustic signal (e.g., prosody (suprasegmental), and speaker identity (indexical) content). The interdependency of these three components raises the possibility that judgments of one component could be influenced through an indirect route of manipulating another component. For example, the segmental component of a persons' speech could be made to sound more or less intelligible by altering the suprasegmental content. Previous studies conducted by Henry, McAuley, and Zaleha (2009), and Melara and Marks (1990) suggest this is a possibility, demonstrating that speech rate is not merely a function of speed within the speech signal, but also is dependent on other interrelated factors. Bergmann (1986) also stated that often segmental irregularities are perceived rather as prosodic (suprasegmental) irregularities and vice versa due to the overlap in components. These findings are consistent with interdependency theories that depict an interdependence of perceptual judgments of pitch, timing, and/or personal profile. More specifically, manipulations of prosodic characteristics, such as fundamental frequency, have been shown to increase perceived fluency in PWS (Wingate, 1969, 1970).

Given that perceptual judgments of prosody may be subject to interdependencies of acoustic factors, a suprasegmental variable worthy of study is the effect of altered F0. In particular, findings of interdependency among components raise the possibility that targeting a property such as overall F0 might influence the perceptions of components that seem to relate more to the segmental component, such as intelligibility or fluency.

The contrasting hypotheses between independent and interdependent speech processing theories were investigated by Goldinger (1998), who put forward an influential episodic model of language processing. Goldinger's episodic model proposed that segmental information, as well as aspects of indexical information shown to be reflected in the speaker's voice, mood, speaking rate, etc., should be explored in combination. The crux of episodic theories is that perceptual details are stored in memory and are integral to later perception (Goldinger,1998, Pierrehumbert, 2003, Pisoni, 1993, Mullennix, Pisoni, & Martin, 1989, Mullennix & Pisoni, 1990). This idea ties into the pre-existing theory of perceptual normalization being an integral component of spoken-word recognition (Mullennix, Pisoni, & Martin, 1989; Mullennix & Pisoni,1990). Although the idea of speaker normalization, as proposed by Halle (1985), and Joos (1948) may be true to some extent, it can be generally stated that normalization is necessary for speech perception to some degree. Many theories of speech perception assume that special processes match stimuli to canonical representations in memory, which is done via speaker normalization. Speaker normalization also states that abstract lexicon can be assumed to be motivated by signal variation and is considered a perceptual problem to be solved by recognition systems of listeners (McClelland & Elman, 1986; Morton, 1969; Studdert Kennedy, 1976, Gertsman, 1968). Additional theories state that automatic normalization plays a significant role, primarily when no effects of speaker variation are evident, implying that an "automatic normalization process occurs"(Goldinger, 1998). Regardless, of differing theories, normalization is reflected in speech perception, but more as a hypothesis, rather than an assumption. However, assumptions can be made regarding memory for stored "surface" details of communication as a whole, including nonlinguistic stimuli (i.e faces, pictures, pitch and tempo, social interactions) (as cited in Goldinger 1998). Additionally, linguistic processes often create lasting, detailed memories (Goldinger, 1998). Semon, 1909, 1923 (as cited in Goldinger, 1998) proposed that words are stored in memory as smaller pieces of a larger context, and without any context, the retrieval seems abstract. Voice information is a critical component of retrieval of context of spoken words Garner (1974).Such episodic and perceptual normalization theories can account for interdependency among segmental, suprasegmental, indexical properties. Recent developments in linguistic theories also support this model, proposing that listeners do in fact retain memory traces of the fine-grained acoustic information (suprasegmental) in the speech signal along with speaker information (indexical) (Goldinger, 1996) This further refutes theories of independence and indicates interdependencies between the various components of segmental, suprasegmental, and indexical aspects of speech. Such a theory could have significant consequences for identification of spoken words, the effects of rate and talker variability, and/or establishing the importance of perceptual normalization (Sommers, Nygaard, & Pisoni, 1994). However, the applicability of such theories as they relate to speech processing issues in clinical speech-language pathology has largely not been investigated.

Due to the suspected interdependencies between speech characteristics, and the limited amount of recent literature, it is critical to continue to investigate the carryover between streams of segmental, suprasegmental, and indexical information within a speech signal. Dilley, Wieland, Gamache, McAuley, and Redford (2013) explored this carryover by manipulating the overall formants and fundamental frequency (F0) of speech for typically-developing five-year old children and adults in order to mimic the indexical, age-related changes in F0 and formants which occur during normal development. Listeners heard speech files in which both the formants and F0 were unmodified (i.e., children sounded like children, and adults sounded like adults), as well as conditions where the formants and F0 had been modified (i.e., the children sounded like adults, and adults sounded like children). Listeners then made judgments of segmental, suprasegmental, and indexical properties of the talker. If segmental and/or suprasegmental properties of speech are processed independently of indexical properties (Joos, 1948; Halle, 1985), then the F0 and formant modification should not have affected the judgments of these properties. Contrary to this view, the results from this study indicated an overlap between the three streams, where the indexical/age-related manipulation affected judgments of segmental and/or suprasegmental properties. Further, these findings indicated that the speech of child talkers, when modified to a lower pitch contour, was judged to be slower, less fluent, and less intelligible. The same child talkers with modified speech were judged to be less likeable, more anxious, and more likely to have cognitive and/or speech-language impairments. The results supported the hypothesis that an interdependence exists between various aspects of the speech signal that can affect a wide range of impressionistic judgments of attributes that are commonly assumed to be orthogonal or quasi-orthogonal to prosodic, segmental, and talker attributes. These findings may be critical to understanding how speech is perceived by a listener, as well as how various factors/streams overlap in order to effectively communicate.

This thesis study builds on the work of Dilley and colleagues (2013) which identified interdependencies in perception among segmental, suprasegmental, and indexical aspects of speech in typically-developing children. Specifically, we investigated whether the interdependencies demonstrated by Dilley et al. (2013) can be generalized to clinical populations such as children who stutter. If this interdependency proves true, the perception of speech would depend on influential factors or patterns of one component by other components (i.e., segmental, suprasegmental, indexical). Therefore, one research question this thesis study will be investigating is whether an F0-only manipulation would provide support for the hypothesis of interdependencies among speech components (segmental, suprasegmental, and indexical) in the clinical population of children who stutter? Another research question that was explored was will manipulating overall F0 to be higher or lower affect impressionistic judgments of segmental and/or suprasegmental properties relating to anxiety, cognitive abilities, fluency, gender, intelligibility, likeability, speech rate, and/or speech-language abilities of children who stutter? This thesis study investigated the effect of one prosodic characteristic (i.e., F0) on other impressionistic judgments of prosodic, segmental and indexical aspects of speech in children who stutter. We also explored whether changing the indexical properties of a speech sample will affect judgments of segmental and/or suprasegmental properties of children who stutter. To test if indexical and/or suprasegmental information would be found to be interdependent of segmental information in speech, we altered the suprasegmental F0 to be higher or lower in pitch, which we expected might be more likely to enhance the segmental perceptions of fluency in a speech sample. Additionally, an F0-only manipulation revealing similar results would demonstrate interdependence among components in perceptual judgments. These findings give rise to the possibility of new clinical approaches aimed at altering volitional F0 in a manner which might result in more positive perceptual evaluations of segmental, suprasegmental, or indexical aspects of speech of PWS.

METHODS

Participants

Thirty three students (F = 25, M = 8) from Michigan State University participated in the experiment in return for partial credit in a university course. Participants were recruited via SONA network, and were native speakers of American English who were at least 18 years of age (M = 21.3) with self-reported normal hearing. The design of the experiment was a 3 (modification condition: Lowered, Original, Raised) x 3 (Severity level: Mild, Moderate, Severe) mixed factorial. Modification and Severity measures were both within-subjects variables.

Stimuli

The stimuli were obtained via recordings of spontaneous speech produced in a child-friendly, quiet, experimental room in the Speech Neurophysiology Laboratory at Michigan State University. Speech samples were collected from 15 children, three to ten-year-old (6 female; 9 male, with a mean age of: 6 months, 2 years, or 74.6 months) who stutter. The 15 children were chosen based on previously established severity of his/her fluency disorder into three groups of Mild, Moderate, and Severe. This was arranged so that five different children, both males and females, were in each of the three severity groups. Severity ratings were determined based on previous diagnoses from speech language pathologists. None of the children had any concomitant speech, hearing, or developmental disabilities, besides fluency.

Speech samples selected for the present study included speech segments extracted from a ten minute narrative task from the wordless picture book “Where are you Frog” by Mercer Meyer (1967). Segments of speech were extracted from the first time the child was exposed to this specific reading stimulus, so that the child had never read/seen the wordless picture book

"Where are you Frog?" prior to this task. First, each full 10-minute recording of the wordless picture book was segmented into three speech segments from each CWS so that each segment contained various disfluencies. Each speech segment ranged from 4-10 seconds long, with the mean duration across all 15 stimuli being 4.2 seconds. This resulted in a total of 45 speech samples (3 segments x 15 children = 45), and approximately 30 seconds of speech from each child (6-10 seconds x 3 segments = approximately 30 seconds) in each severity group prior to manipulations. Disfluencies were characterized by syllable repetitions, whole or part word repetitions, prolongations of audible or inaudible speech sounds, and/or unintentional delays in initiating sound (ASHA, 2004). Other disfluencies (OD's) that were included in the segmentation included interjections, revisions, and phrase repetitions.

Spectral Modifications of Stimuli

Spectral modifications to alter F0 of each of the three speech segments from each CWS was conducted in Praat using spectral modification tools (Boersma & Weenink, 2002). Specific values for global settings for the pitch measurement factors and modification parameters were restricted to particular ranges for natural-sounding child speech. The first step to completing spectral modifications using Praat was to change global spectral parameters for a talker using pitch-synchronous overlap-add algorithm (Moulines & Charpentier, 1990). This involved setting a global value for each talker for two pitch measurement factors (i.e., pitch floor, and pitch ceiling), and two modification parameters (i.e., formant shift ration and new pitch median). The standard values for child speech were as follows:

A. Time step (s) = 0.01
B. Minimum pitch (Hz) = 120 Hz for children

C. Maximum Pitch (Hz) = 700 Hz for children

Two other modification parameters (i.e., pitch range factor and duration factor) were left at the default value of 1.0 for all talkers. Next, the three extracted speech segments from the 15 CWS, each underwent separate modifications of F0 values using Praat by shifting the frequencies of the F0 from the original sound file. Values for shifting frequencies were selected for which the talker seemed to be most convincingly transformed in pitch, yet still natural sounding for the child's chronological age and gender. For each segment, the frequencies were shifted by -50.0, 0.0, and +50.0 Hz to produce the modified Lowered, Original, and Raised speech segments, respectively. Manipulation for the Original level assured that artifacts created by manipulation itself existed in all three levels, not just the speech samples where the F0 frequency was altered. Fifty hertz was chosen as the amount of modification because it corresponded to 1.32 standard deviations for children ages 5-10 as reported in Lee et al., 1999, JASA paper on children's speech qualities. These modifications allowed for holding constant the prosodic attributes, such as speech rate, F0 contour, pause timing, and duration. The modifications resulted in nine speech files per CWS (3 speech segments (1, 2, 3) x 3 modification levels (Lowered, Original, Raised). Therefore, there were 135 total speech segments across all CWS, so that all 3 speech segments, and all 3 modification conditions were represented for all 15 CWS (3 modification types x 3 modification levels x 15 =135). The Lowered, Original, and Raised modification conditions were expected to correspond to speakers' sounding older, of the same age, or younger, respectively.

Following the modifications, the speech sounded very natural for the age of the children, with little if any perceptual indication of modification. To further minimize possible artifacts of modification, the speech was carefully checked auditorily by researchers and piloted. Additionally, participants were asked to complete a post-experiment questionnaire in which they identified any patterns or abnormalities in the speech sample in order to further asses any perceptions of oddities observed in the speech samples. Of 33 participants, only 5 participants (15%) made any mention to observable oddities throughout the experiment. Some of the oddities described included poor sound quality, background noise, and computerized voices.

Stimulus Lists

After the modifications were completed, three stimulus lists were created consisting of 5 practice trials and 45 experimental trials. The 5 practice trials oriented participants to the task and allowed them the opportunity to ask questions. Speech segments used in the practice trials came from different CWS who were not represented in the experimental trials. E ach of the three conditions and severity levels were represented in the practice block. The 45 experimental trials consisted of all three speech segments from each child (15 children who stutter), with each segment represented in a different modification condition (Lowered, Original, Raised) (15 x 3 = 45). The experimental speech segments were then divided into 3 blocks of 15 samples with a mandatory short break of 45 seconds between each block. Three blocks were utilized in order to break up the speech samples for two main reasons. The first reason was in order to give the listener a short break. The second reason was in order to present a new speech segment from each of the CWS in a new condition (Original, Raised, Lowered).

To create lists for the experiment, stimuli were first blocked by child speaker, so all listeners heard and judged one fragment from each of the 15 talkers before hearing additional fragments from a speaker. The order of presentation of child speakers within blocks was pseudorandomized, with the constraint that CWS within the same severity level were not heard consecutively. The same sequence of child talkers was used for all blocks. Three stimulus lists were then created using a Latin Square experimental design by counterbalancing the pairing of modification condition (Raised, Original, Lowered) with speech fragments. A Latin Square design ensured that each listener heard each of the three speech samples from each of the 15 CWS only once throughout the experiment while being exposed to all Modification conditions. The pairing of Modification conditions with speech fragments in each list was further constrained within a block such that listeners heard a specific sequence of Modification conditions that repeated every three stimuli within a list (e.g., Original, Raised, Lowered, Original, Raised, Lowered, etc. for the first block).

The stimulus lists are represented in Appendix A: Table 1: Stimulus Lists.

Task

Participants listened to 50 short speech fragments over headphones and answered 9 questions immediately following each speech fragment. The questions that followed asked the participant to provide impressionistic judgements of the speaker on nine attributes of the speaker based upon an eight point scale, which was displayed on the screen.

The first rating that the participants were asked to provide impressionistic judgments on was in regards to the age of the speaker (0-100 years). Participants estimated the speaker's age in years and months using the format Age. Month (age separated from month by a period). For example, 1 year and 3 months was depicted as 1.3. Participants were told in instructions that typically developing children begin speaking by one year of age and begin speaking in phrases by three years of age.

The second rating that the participant was asked to provide impressionistic judgments on was in regards to the gender of the speaker. Participants assessed his/her confidence in regards to the gender of the speaker, with 1 indicating that he/she was positive the speaker was a male, and 8 indicating that he/she was positive the speaker was a female.

The third rating that the participants were asked to provide was in regards to the speech rate of the speaker. Participants assessed his/her perception of how fast the speaker was speaking, with 1 indicating that the speaker was speaking fast, and 8 indicating that the speaker was speaking slow.

The fourth rating that the participants were asked to provide was in regards to the intelligibility of the speaker. Participants assessed his/her perception of what percentage of words were understood, with 0 indicating 0%, or none of the words were able to be understood, and 100 indicating that 100%, or all of the words were able to be understood. Participants were informed that intelligibility refers specifically to how much you could understand from what they were actually saying.

The fifth rating the participants were asked to provide was in regards to the speaker's cognitive abilities. Participants assessed his/her confidence in regards to the overall cognitive, or mental abilities, of the speaker, with 1 indicating that he/she was positive that the speaker had no cognitive impairment, and 8 indicating that he/she was positive the speaker did have a cognitive impairment. Participant were informed that cognitive abilities refers specifically to the underlying mental abilities of the speaker, and they were asked to rate the speaker on whether or not you believe there could be some sort of mental delay or disability.

The sixth rating the participants were asked to provide was in regards to the speaker's speech and language abilities. Participants assessed his/her confidence in regards to the overall speech and language abilities of the speaker, with 1 indicating that he/she was positive that the speaker had no speech-language delay or disorder, and 8 indicating that he/she was positive that the speaker did have a speech-language delay or disorder. Participants were informed that the speech and language delay/disorder rating differs from the cognitive abilities rating in that the speech and language rating refers specifically to the output of speech, rather than the underlying abilities. Participants were reminded to remember to include his/her perception of both disorders and delays in the speaker.

The seventh rating the participants were asked to provide was in regards to the speaker's anxiety. Participants assessed his/her confidence in regards to the overall anxiety of the speaker, with 1 indicating that the speaker was completely non-anxious/comfortable, and 8 indicating that the speaker was completely anxious/uncomfortable. Participants were informed that anxiety could be evidenced by how timid or shy the speaker was, or how excitable the speaker was in the speaking situation.

The eighth rating the participants were asked to provide was in regards to the speaker's likeability. Participants assessed his/her confidence in regards to the overall likeability of the speaker, or how enjoyable he/she would be as a conversational partner, with 1 indicating that the speaker was strongly likeable, and 8 indicating that the speaker was strongly unlikeable.

The ninth, and final rating that the participants were asked to provide was in regards to the speaker's fluency. Participant assessed the overall degree of fluency of the speaker when speaking, with 1 indicating that the speaker was completely fluent, and 8 indicating that the speaker was completely disfluent. Participants were informed that the term fluency in this case referred to the broad spectrum of fluent speech characteristics, whether by normal talkers or people who stutter, and that disfluency referred to the broad spectrum of disfluent speech characteristics. Participants were directed to pay attention to the difference between fluency and intelligibility ratings, Intelligibility is again referring to how much you can understand, whereas fluency is referring to the overall fluidity of the speech. Therefore, it is possible for an individual to be completely intelligible, while simultaneously being completely disfluent. Participants were given a reference sheet hat outlined each of the attributes and the rating scales for reference throughout the experiment.

Procedure

The 33 participants were assigned randomly to each of the three stimulus lists (n = 11 per list). Before each experiment, participants completed 5 practice trials, and then given a chance to ask questions before beginning the actual experiment. Participants were not made aware of the speech modifications. Throughout the experiment, the scale associated with each question was displayed on the screen in order to reinforce consistent use of scale endpoint referents, as well as to encourage use of intermediate variables.

Apparatus

Participants listened to the speech fragments over Sennheiser HD 280 headphones. The experiment was presented to participants using E-Prime Version 2.0 Professional (Psychology Software Tools, Inc.) running on a Lenovo Intel Core2 Duo CPU E8500 computer with a 19-in. monitor.

RESULTS

A 3x3 repeated measures ANOVA with factors of modification (Lowered, Original, Raised conditions), and severity level (Mild, Moderate, Severe) was conducted for each dependent variable. Results for these dependent variables are discussed in turn below.

Age

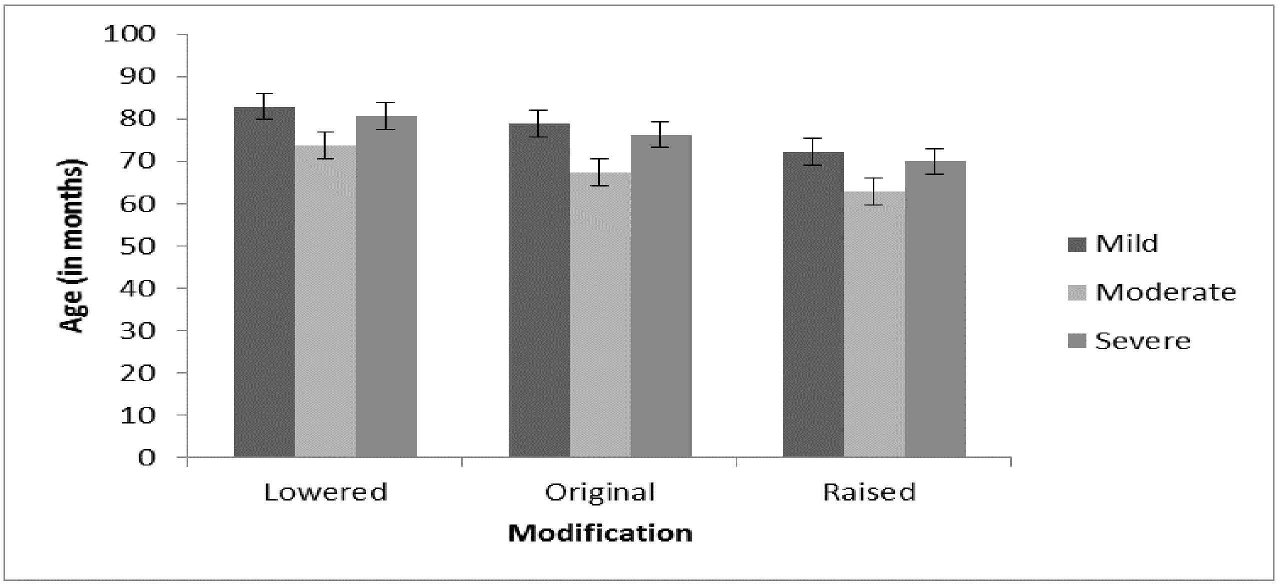

Figure 1. Age ratings as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). Participants were asked to estimate the speaker's age in years and months. The average age in months is represented in the graph above. Error bars indicate ± 1 SEM.

Due to a coding glitch, the deliminator that was used to separate years from months was not recorded in the program throughout the experiment, yielding a concatenation of years and months in digits in which the initial zeros were omitted. Data was therefore hand coded to separate digit strings into years and months. In doing so, experimenter assumptions had to be made. Reasonable assumptions were based upon the fact that participants were informed upon initial instruction that children typically begin speaking first words at approximately 1 year of age, and begin speaking in phrases at about 3 years of age. Single digits 1-9 were thus converted to represent a round number of years, assuming that participants would not estimate a child speaking in phrases to be less than 12 months. For responses that indicated 10-16, the initial two numbers were converted to represent years, with the final digit assumed to represent months. . For example, if a response was noted as 111, it was assumed that the participant intended the string of digits to represent 11.1, eleven years, one month. Throughout the analyzing of this data, intent assumptions were made consistently across all modifications and severity levels alike, therefore limiting possible confounds for the variables in questions.

Figure 1 shows the average rating for perceived age in months for each modification and severity condition. There was a significant main effect of modification on ratings of age ($F(2,64) = 63.47$, $p < .001$, $\eta_p^2 = .672$). Moreover, there was a significant main effect of severity ($F(2,64) = 25.316$, $p < .001$, $\eta_p^2 = .450$). Finally, there was no significant interaction between modification and severity ($F(4,124) = .248$, $p = .910$, $\eta_p^2 = .008$).

Next, a one-way ANOVA was conducted on modification. There was a significant effect of modification ($F(2,64) = 63.47$, $p < .001$, $\eta_p^2 = .672$). Post hoc, Bonferroni-corrected pairwise t-test comparisons demonstrated that CWS in the Original condition ($M = 73.51$ months) were perceived as being older than CWS in the Raised condition ($M = 67.89$ months, $p < .001$), and as being younger than CWS in the Lowered ($M = 78.69$, p < .001) modification conditions. Additionally, CWS in the Lowered modification condition were perceived as being significantly older than CWS in the Raised modification condition ($p <.001$).

Finally, a one-way ANOVA was conducted on severity. There was a significant effect of severity ($F(2,64) = 25.316$, $p < .001$, $\eta_p^2 = .450$). Post hoc, Bonferroni-corrected pairwise t-test comparisons demonstrated that children in the Mild severity condition ($M = 77.32$ months) were perceived as being older than children in both the Moderate (*M* = 67.69 months, $p < .001$), and CWS in the Severe (*M* = 75.08, $p < .001$). Additionally, CWS in the Severe severity group were perceived as being older than CWS in the Moderate ($p < .001$) severity conditions.

There was no significant interaction between modification and severity ($F(4,124) = .248$, $p = .910$, $\eta_p^2 = .008$).

Anxiety

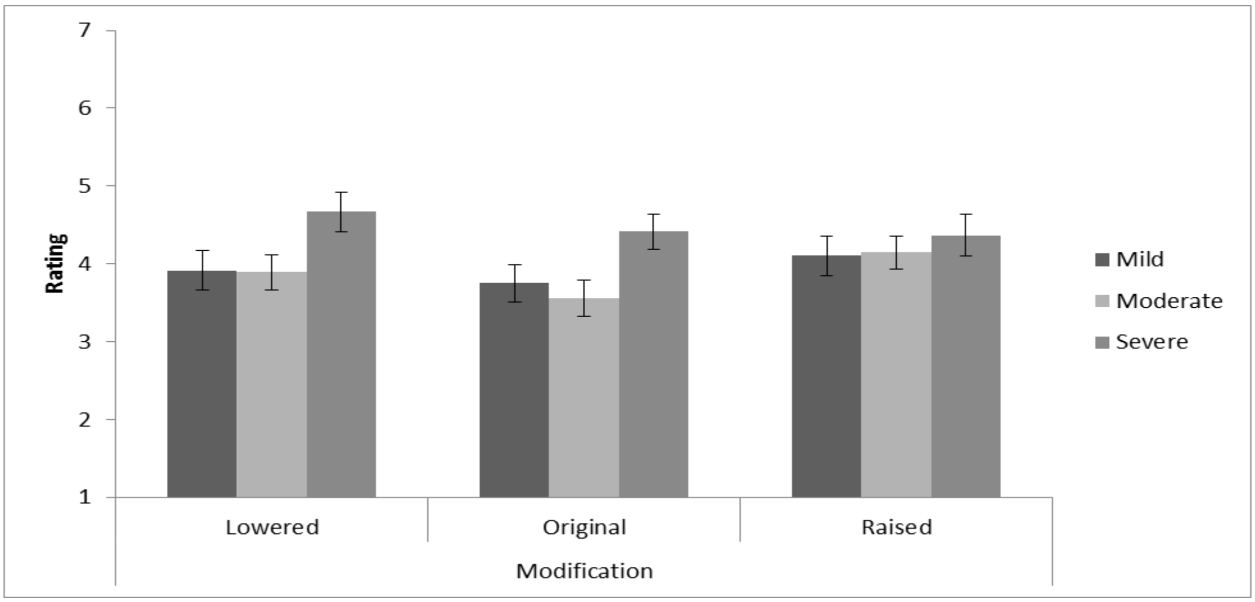

Figure 2. Anxiety ratings as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). Ratings were assessed with 1 indicating listener was positive the speaker was non-anxious and comfortable in the speaking situation, and 8 indicating the listener was positive the speaker was anxious and uncomfortable in the speaking situation. Error bars indicate ± 1 SEM.

Figure 2 shows the average rating for anxiety for each modification and severity condition. There was a significant main effect of modification on ratings of anxiety ($F(2,64) = 5.029$, $p = .009$, $\eta_p^2 = .136$). Moreover, there was a significant main effect of severity ($F(2,64) =$

24.058, $p < .001$, $\eta_p^2 = .429$). Finally, there was a significant interaction between modification and severity ($F(4,128) = 2.622$, $p = .038$, $\eta_p^2 = .076$).

In order to tease apart the significant interaction, a one-way ANOVA was conducted at each level of severity to examine the simple effects of modification. The only level of severity that showed a significant effect of modification was the Moderate level ($F(2,64) = 6.853$, $p = .003$, $\eta_p^2 = .171$). For the Moderate level, post-hoc, Bonferroni-corrected, t-tests showed the Original modification ($M = 3.56$) was perceived as significantly lower in anxiety than the Raised ($M = 4.15$, $p = .005$).

Cognitive Abilities:

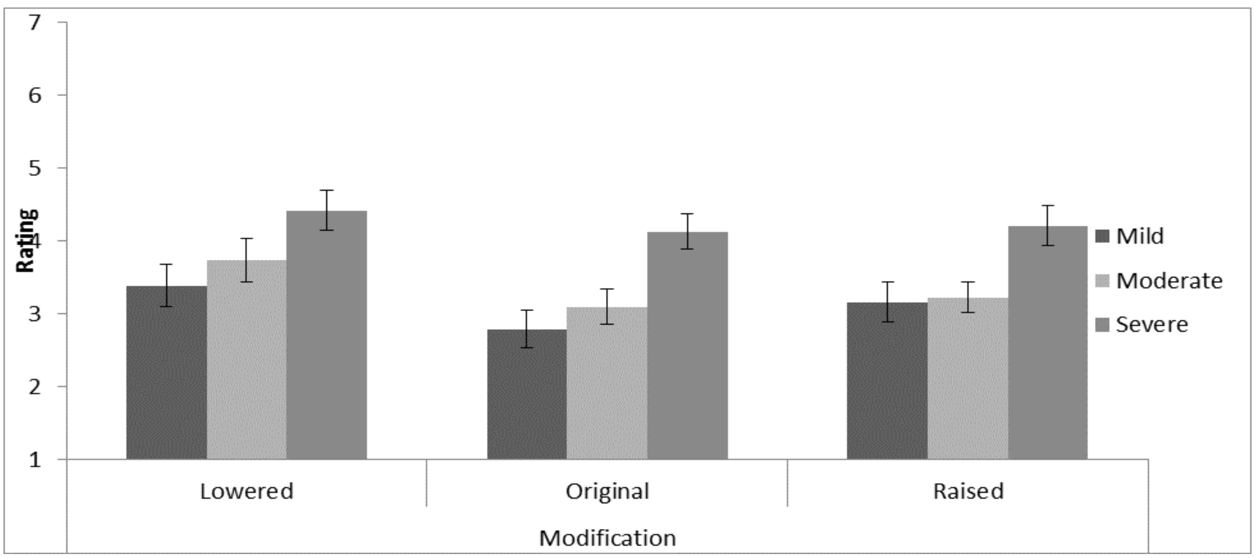

Figure 3. Cognitive ability ratings as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). Ratings were assessed with 1 indicating listener was positive the speaker had no cognitive impairment, and 8 indicating the listener was positive the speaker did have a cognitive impairment. Error bars indicate ± 1 SEM.

Figure 3 shows the average rating for perceived cognitive abilities for each modification and severity condition. There was a significant main effect of modification on ratings of cognitive ability ($F(2,64) = 12.420$, $p < .001$, $\eta_p^2 = .280$). Moreover, there was a significant main effect of severity ($F(2,64) = 24.058$, $p < .001$, $\eta_p^2 =.657$). Finally, there was no significant interaction between modification and severity ($F(4,128) = 2.622$, $p = .038$, $\eta_p^2 =.076$).

Next, a one-way ANOVA was conducted on modification. There was a significant effect of modification ($F(2,64) = 12.420$, $p < .001$, $\eta_p^2 = .280$). Post hoc, Bonferroni-corrected pairwise t-test comparisons demonstrated that speakers in the Original condition ($M = 3.34$) were perceived as having significantly better cognitive abilities than in the Raised condition ($M = 3.53$, $p < .014$), and Lowered ($M = 3.85$, $p < .001$) modification conditions.

Finally, a one-way ANOVA was conducted on severity. There was a significant effect of severity ($F(2,64) = 61.210$, $p < .001$, np2 = .657). Post hoc, Bonferroni-corrected pairwise t-test comparisons demonstrated that children in the Mild severity condition ($M = 3.111$) were perceived as having significantly better cognitive abilities than children in the Moderate ($M = 3.352$, $p = .033$) and in the Severe ($M = 4.251$, $p < .001$) severity conditions. Additionally, children in the Moderate severity condition were perceived as having significantly better cognitive abilities than in the Severe severity condition ($p < .001$).

Fluency

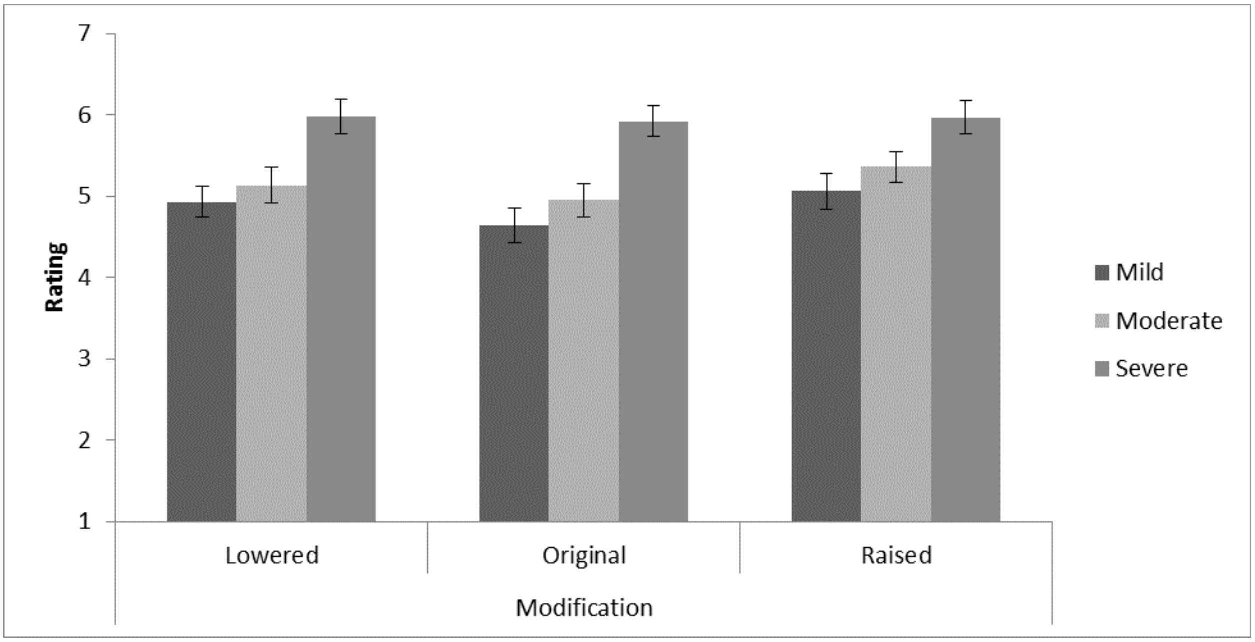

Figure 4. Fluency ratings as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). Ratings were assessed with 1 indicating listener perceived the speaker to be completely fluent, and 8 indicating the listener perceived the speaker to be completely disfluent. Error bars indicate ± 1 SEM.

Figure 4 shows the average rating for fluency for each modification and severity condition. There was a significant main effect of modification on ratings of fluency ($F(2,64) = 5.701$, $p = .005$; $\eta_p^2 = .151$). Moreover, there was a significant main effect of severity ($F(2,64) = 55.249$, $p < .001$; $\eta_p^2 < .001$). Finally, there was no significant interaction found between modification and severity ($F(4,128) = 0.955$, $p = .435$; $\eta_p^2 = .029$).

Next, a one-way ANOVA was conducted on modification. There was a significant effect of modification ($F(2,64) = 5.701$, $p = .005$; $\eta_p^2 = .151$). Post hoc Bonferroini corrected pairwise comparisons demonstrated that the Original modification ($M = 5.170$) was significantly more fluent than the Raised ($M = 5.463$, $p = .016$) modification condition.

Finally, a one way ANOVA was conducted on severity. There was a significant main effect of severity observed ($F(2,64) = 55.249$, $p < .001$; $\eta_p^2 < .001$). Post hoc Bonferroini corrected pairwise comparisons demonstrated that the Mild ($M = 4.875$) was perceived as significantly more fluent than the Moderate ($M = 5.147$, $p = .014$), and the Severe ($M = 5.956$, $p < .001$) severity levels. Additionally, the Moderate severity level was perceived as significantly more fluent than the Severe ($p < .001$) severity level.

No significant interaction was found between modification and severity ($F(4,128) = 0.955$, $p = .435$; $\eta_p^2 = .029$).

Gender

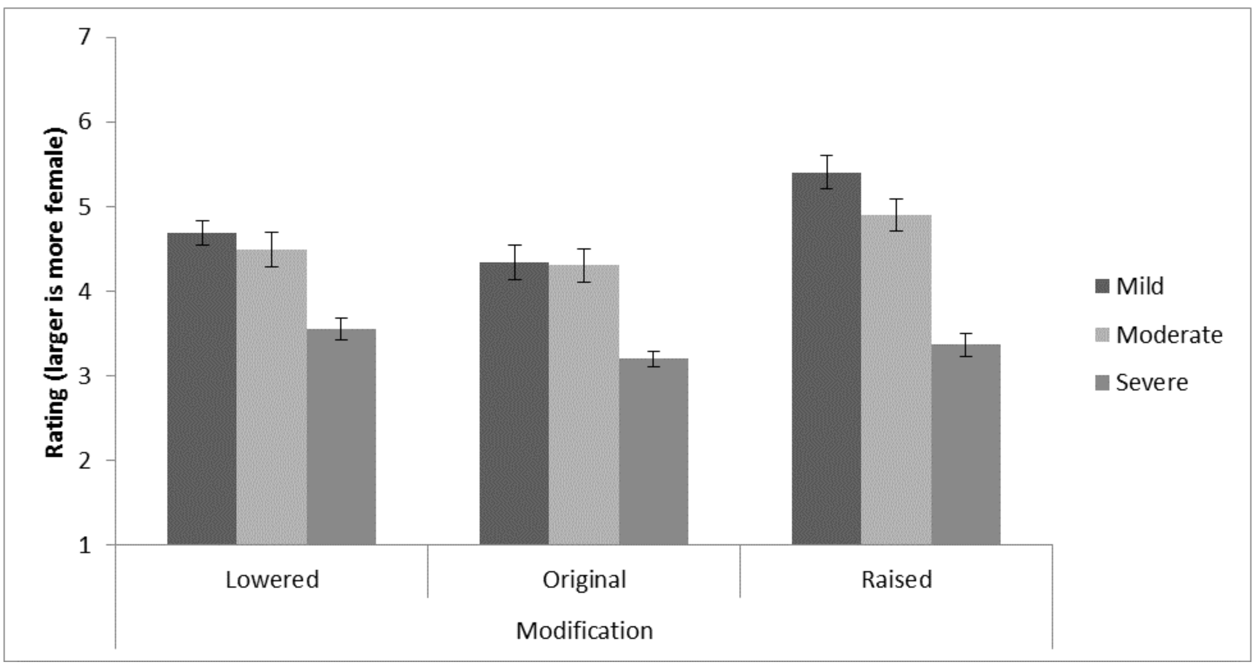

Figure 5. Gender ratings as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). Ratings were assessed with 1 indicating listener perceived the speaker to be positively male, and 8 indicating the listener perceived the speaker to be positively female. Error bars indicate ± 1 SEM.

Figure 5 shows the average rating for gender for each modification and severity condition. A significant main effect of modification was observed when collapsing across all severity levels ($F(2,64) = 11.173$; $p < .001$; $\eta_p^2 =.259$). Moreover, there was a significant main effect of severity observed ($F(2,64) = 66.315$, $p < .001$, $\eta_p^2 = .675$). Finally, there was a significant interaction found between modification and severity ($F(4,128) = 4.319$, $p = .003$, $\eta_p^2 = .119$).

A one-way ANOVA was conducted on modification. There was a significant main effect of modification ($F(2,64) = 11.173$; $p < .001$; $\eta_p^2 =.259$). Post-hoc Bonferroni-corrected pairwise comparisons demonstrated that the Original modification ($M = 3.95$) was perceived as significantly less female than the Raised ($M = 4.56$, $p < .001$) and trended towards less female for the Lowered ($M = 4.24$; $p =.077$) modifications.

Finally, a one-way ANOVA was conducted on severity. There was significant main effect of severity was observed when collapsing across all modification levels ($F(2,64) = 66.315$, $p < .001$, $\eta_p^2 = .675$). Post-hoc Bonferroni-corrected pairwise comparisons demonstrated that the Severe severity ($M = 3.37$) was perceived as significantly less female than the Moderate ($M = 4.57$, $p < .001$) and Mild ($M = 4.81$, $p < .001$) severity levels.

A significant interaction was found between modification and severity ($F(4,128) = 4.319$, $p = .003$, $\eta_p^2 = .119$). Therefore, a one-way ANOVA was conducted to look at the simple effects of modification at each level of severity. The Mild severity showed a significant effect of modification ($F(2,64) = 13.200$, $p < .001$, $\eta_p^2 = .292$), and Bonferroini-corrected post-hoc tests showed the Raised modification ($M = 5.41$) was perceived as significantly more female than the Original ($M = 4.34$, $p < .001$), and the Lowered ($M = 4.69$, $p = .005$) modification conditions. The Moderate severity showed a significant effect of modification ($F(2,64) = 4.267$, $p = .018$, η_p^2

= .118), and Bonferroini-corrected post-hoc tests showed that the only significant difference was where Original (*M* = 4.30) was perceived as significantly less female Raised (*M* =4.90, *p* = .031) modification conditions. The Severe severity had no significant main effect of modification.

Intelligibility

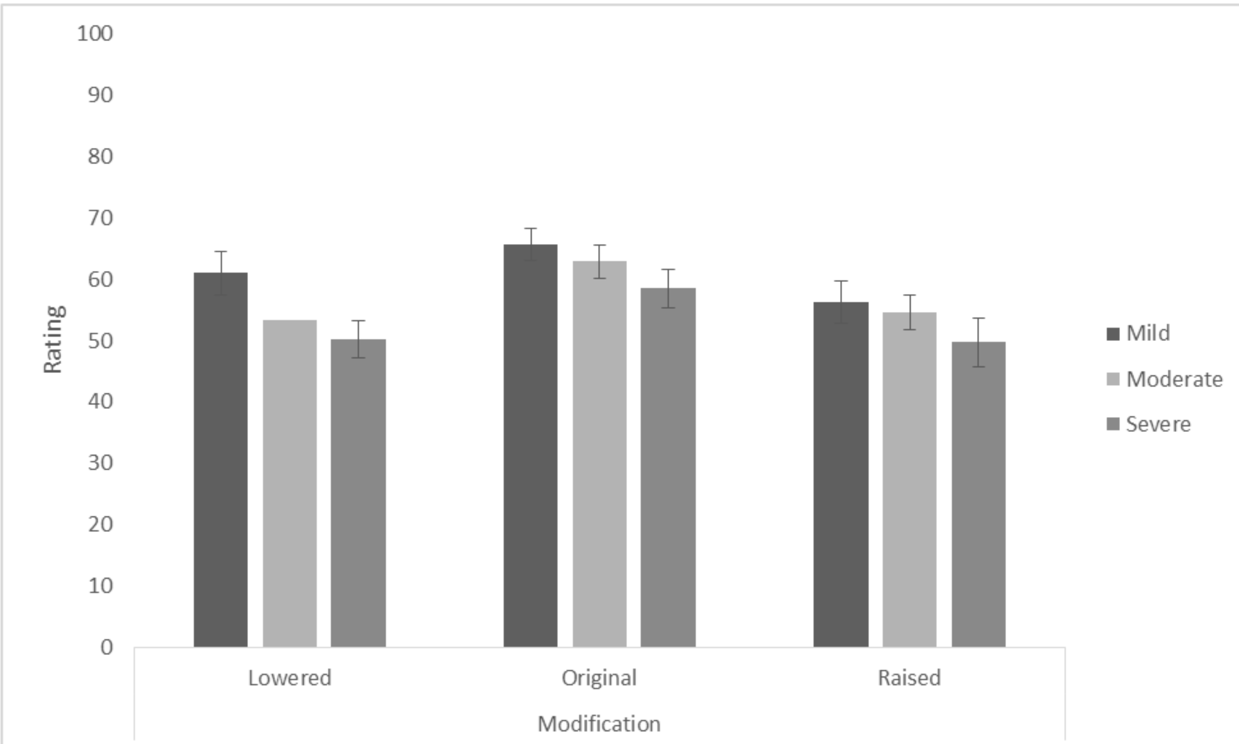

Figure 6. Percentage of intelligible words as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). This dependent variable was assessed by listeners providing a perceived percentage of words understood between 0%, indicating none of the words were understood, and 100%, indicating that all of the words were understood. Error bars indicate ± 1 SEM.

Figure 6 shows the average percentage of perceived intelligible words for each modification and severity condition. A significant main effect of modification was observed when collapsing across all severity levels ($F(2,64) = 11.530$, $p < .001$, $\eta_p^2 = .265$). Moreover, there was a significant main effect of severity ($F(2,64) = 17.481$, $p < .001$, $\eta_p^2 = .353$). Finally, there was no significant interaction was found between modification and severity with ($F(4,128)$ $=.612$, $p = .655$; $\eta_p^2 = .019$).

A one-way ANOVA was conducted on modification. There was a significant effect of modification ($F(2,64) = 11.530$, $p < .001$, $\eta_p^2 = .265$). Post-hoc Bonferroni-corrected pairwise comparisons demonstrated that the Original modification ($M = 62.33$) was perceived as significantly more intelligible than the Lowered ($M = 54.75$; $p = .003$) and the Raised ($M = 53.53$, $p < .001$) modification conditions.

Finally, a one-way ANOVA was conducted on severity. There was a significant main effect of severity was observed when collapsing across all modification levels ($F(2,64) = 17.481$, $p < .001$, $\eta_p^2 = .353$). Post-hoc Bonferroni-corrected pairwise comparisons demonstrated that the Mild severity ($M = 60.98$) was perceived as significantly more intelligible than the moderate ($M = 56.91$, $p = .007$) and severe ($M = 52.71$, $p < .001$) severity levels. Additionally the Moderate severity level was perceived as significantly more intelligible than the Severe ($p = .024$) severity level.

No significant interaction was found between modification and severity ($F(4,128) =.612$, $p = .655$; $\eta_p^2 = .019$).

Likeability

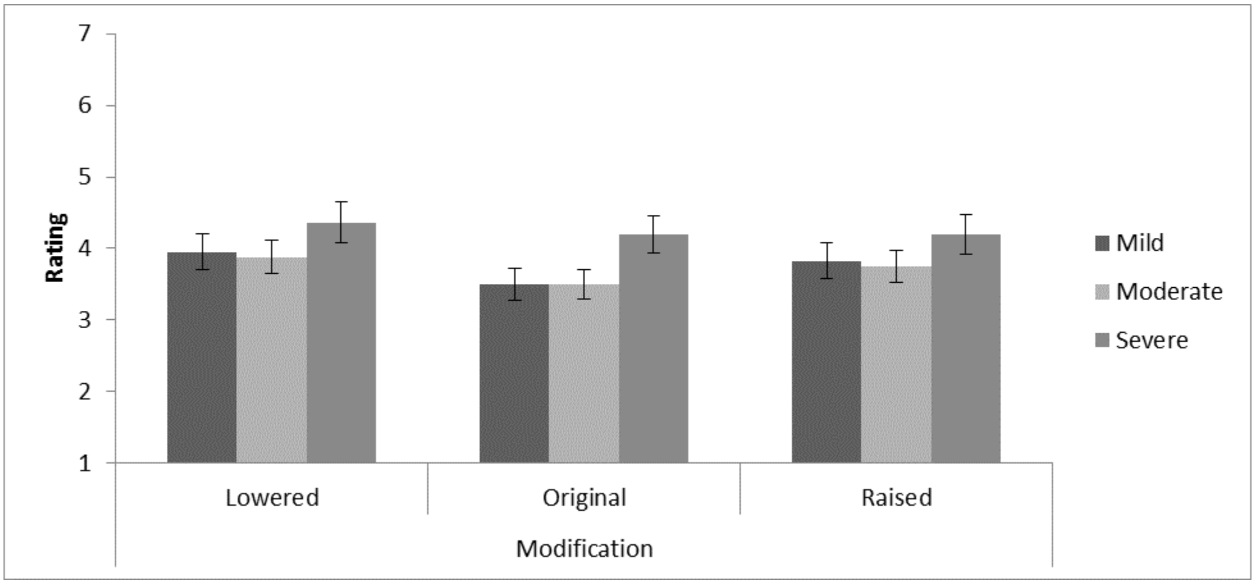

Figure 7. Likeability ratings as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). Ratings were assessed with 1 indicating listener perceived the speaker to be strongly likeable, and 8 indicating the listener perceived the speaker to be strongly unlikable. Error bars indicate ± 1 SEM.

Figure 7 shows the average rating for each modification and severity condition. A significant main effect of modification was observed when collapsing across all severity levels ($F(2,64) = 9.434$; $p < .001$; $\eta_p^2 = .228$). Moreover, a significant main effect of severity was observed ($F(2,64) = 13.450$; $p < .001$; $\eta_p^2 = .296$). Finally, there was no significant interaction was found between modification and severity ($F(4,128) = .827$, $p = .510$; $\eta_p^2 = .025$).

A one-way ANOVA was conducted on modification. There was a significant effect of modification ($F(2,64) = 9.434$; $p < .001$; $\eta_p^2 = .228$). Post hoc Bonferroini corrected pairwise comparisons demonstrated that the Original modification ($M = 3.729$) was perceived as significantly more likable than the Lowered ($M = 4.065$; $p = .001$), and the Raised ($M = 3.923$, $p = .046$) modification conditions.

A one-way ANOVA was conducted on severity. There was significant main effect of severity was observed ($F(2,64) = 13.450$; $p < .001$; $\eta_p^2 = .296$). Post hoc Bonferroini corrected pairwise comparisons demonstrated that the Severe severity condition ($M = 4.248$) was perceived as significantly less likeable than the Mild ($M = 3.758$, $p < .001$) and the Moderate ($M = 3.711$; $p < .001$) severity levels.

No significant interaction was found between modification and severity ($F(4,128) = .827$, $p = .510$; $\eta_p^2 = .025$).

SpeechRate

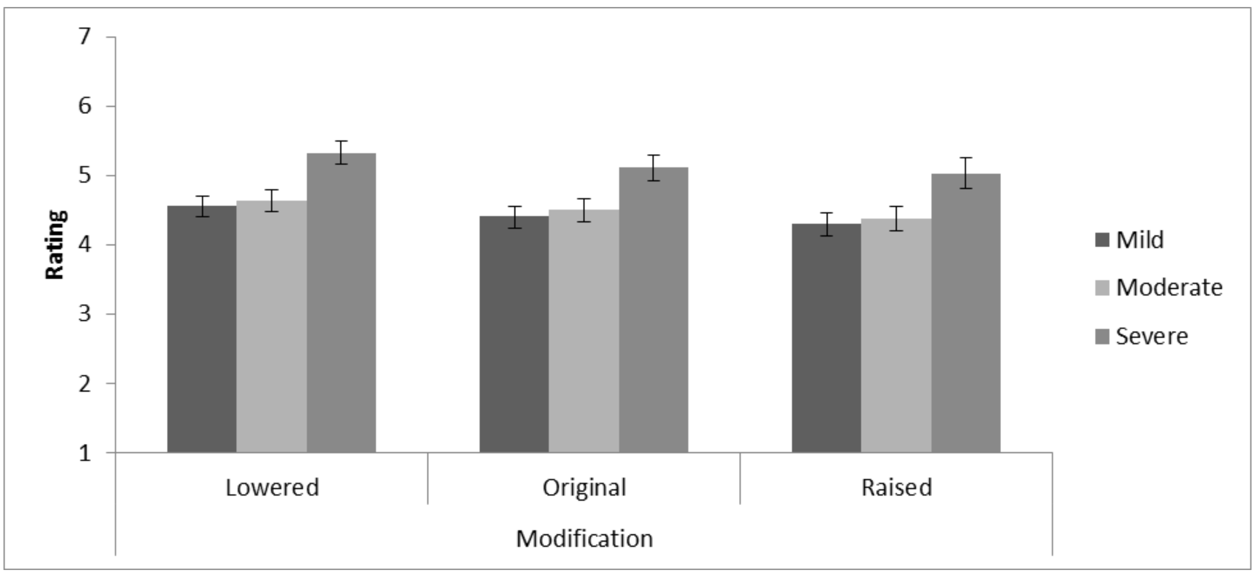

Figure 8. Rate of speech ratings as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). Ratings were assessed with 1 indicating listener perceived the speaker to be speaking fast, and 8 indicating the listener perceived the speaker to be speaking slowly. Error bars indicate ± 1 SEM.

Figure 8 shows the average rating for speech rate for each modification and severity condition. No significant main effect of modification was observed when collapsing across all severity levels ($F(2,64) = 2.794$; $p = .069$; $\eta_p^2 = .080$). However, a significant main effect of severity was observed ($F(2,64) = 37.754$; $p < .001$; $\eta_p^2 = .541$). There was no significant interaction found between modification and severity ($F(4,128) = 0.062$, $p = .993$; $\eta_p^2 = .002$).

A significant main effect of severity was observed when collapsing across all modification levels ($F(2,64) = 37.754$; $p < .001$; $\eta_p^2 = .541$). Post hoc Bonferroini corrected pairwise comparisons demonstrated that the Mild ($M = 4.422$) was perceived as significantly lower than the Severe ($M = 5.156$; $p < .001$) modification condition, and that Moderate severity condition ($M = 4.501$) was significantly lower than the Severe severity condition ($p <.001$).

No significant interaction was found between modification and severity ($F(4,128) = 0.062$, $p = .993$; $\eta_p^2 = .002$).

Speech and Language Abilities

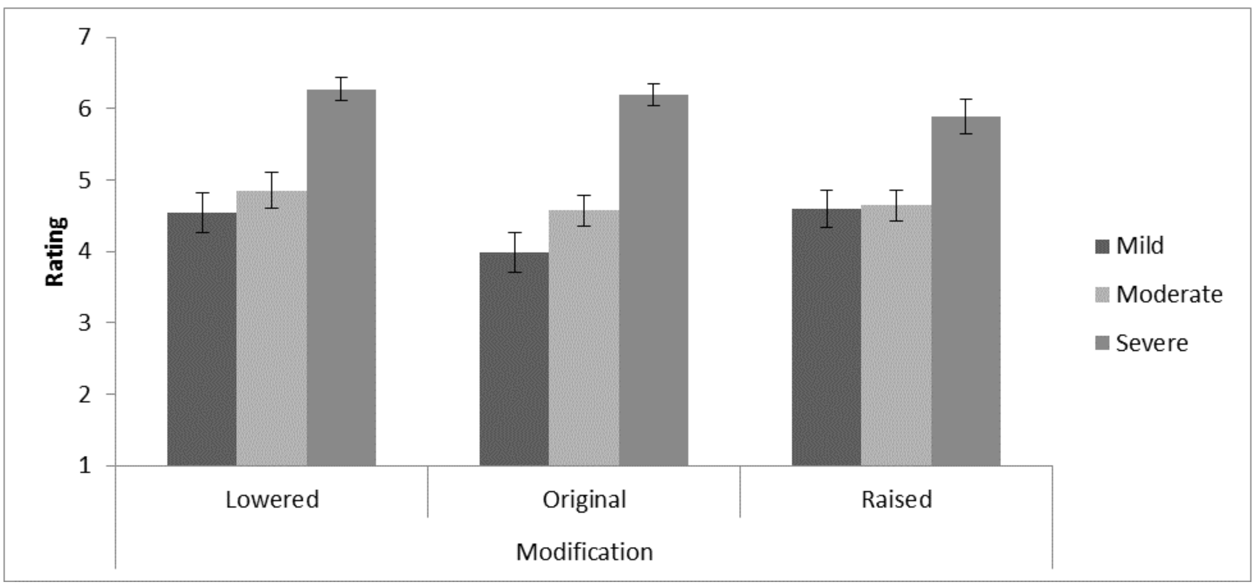

Figure 9. Speech and language ability ratings as a function of modification condition (Lowered, Original, Raised), and severity level (Mild, Moderate, Severe). Ratings were assessed with 1 indicating listener was positive the speaker had no speech and/ or language delay and/or disorder, and 8 indicating the listener was positive the speaker did have a speech and/or language delay and/or disorder. Error bars indicate ± 1 SEM.

Figure 9 shows the mean rating for speech and language abilities for each modification and severity condition. A significant main effect of modification was observed ($F(2,64) = 5.936$, $p = .004$; $\eta_p^2 = .156$). Moreover, a significant main effect of severity was observed ($F(2,64) = 86.447$, $p < .001$, $\eta_p^2 = .730$). Finally, a significant interaction was found between modification and severity ($F(4,128) = 3.536$, $p = .009$; $\eta_p^2 = .100$).

Next, a one-way ANOVA was conducted on modification. There was a significant main effect of modification ($F(2,64) = 5.936$, $p = .004$; $\eta_p^2 = .156$). Post hoc Bonferroini corrected pairwise comparisons demonstrated that the Original modification ($M = 4.919$) was perceived as significantly lower than the Lowered modification ($M = 5.222$; $p = .014$), and the Raised modification ($M = 5.042$, $p < .001$) levels.

A one-way ANOVA was conducted on severity. There was a significant main effect of severity was observed ($F(2,64) = 86.447$, $p < .001$, $\eta_p^2 = .730$). Post hoc Bonferroini corrected pairwise comparisons demonstrated that the Mild severity ($M = 4.374$) was perceived as having significantly better speech and language skills than the Moderate ($M = 4.691$, $p = .055$) and the Severe ($M = 6.119$, $p < .001$) severity conditions. Additionally, the Moderate severity was perceived as having significantly better speech and language skills than the Severe severity level ($p < .001$).

A significant interaction was found between modification and severity ($F(4,128) = 3.536$, $p = .009$; $\eta_p^2 = .100$). Therefore, a one-way ANOVA was conducted to look at the simple effects of modification at each level of severity. The only level of severity that showed a significant effect of modification was the Mild severity level ($F(2,64) = 6.853$, $p = .003$, $\eta_p^2 = .171$), and Bonferroini-corrected post-hoc tests showed the Original modification ($M = 3.99$) was perceived as having significantly better speech and language abilities than the Raised (M = 4.59, p = .001) and trended towards being better for the Lowered (M = 4.54, p = .043) severity level.

DISCUSSION

In this experiment, spectral modification was used to alter fundamental frequency (F0) characteristics for young children who stutter (CWS). This was done in order to determine the possible effect that such modifications may have on the impressionistic judgments of suprasegmental, segmental, and indexical characteristics of a speaker, particularly regarding the clinical population of children who stutter with varying stuttering severity levels. Naïve listeners judged modified speech segments and provided impressionistic judgements on perceived age, anxiety levels, cognitive abilities, gender, intelligibility, likeability, speech rate, and speech and language abilities. Spectral modification was used to allow for targeted manipulation of F0 to examine its effects on impressionistic judgments while controlling for other acoustic variables, including duration, speech rate, overall F0 contour, and pausing, as well as factors such as linguistic content and voice quality.

A main finding of this research was that modifications of F0 influenced naive listener perceptions on all measures examined. The selected measures related to a variety of suprasegmental, segmental, and indexical characteristics for this targeted clinical population of CWS, including speech quality (i.e., intelligibility and fluency). In the following, we will recap the main findings, describing in turn how F0 modifications and severity influenced impressionistic judgments for CWS. We then describe the theoretical and clinical implications of this work and identify some future directions for the research.

Effects of F0 modifications

Segmental and suprasegmental attributes for which F0 significantly affected judgments included fluency, intelligibility, and speech rate. CWS were perceived overall as being significantly more fluent in the Original modification condition than in the raised modification condition. Additionally, CWS were perceived to be significantly more intelligible in the Original modification condition than in the Lowered or the Raised modification. Contrary to ratings of fluency and intelligibility, there was no significant interaction of modification on F0 regarding the suprasegmental attribute of speech rate. However, the *p*-value (p = .069) may indicate a possibility of a Type Two error, thereby implying an effect that could not be detected reliably in the present study design. Overall, these results imply that F0, a primarily suprasegmental characteristic, can influence judgements of other suprasegmental and segmental attributes of CWS. The fact that many of the impressionistic judgements were often more positive in the Original modification condition may suggest that listeners may have been sensitive to whether F0 levels were age-appropriate for the CWS, such that having an age-appropriate F0 may be important for optimal perceptual evaluations of fluency and intelligibility.

F0 modification was also shown to significantly influence perceptual evaluations of indexical attributes for CWS. Recall that indexical attributes refer to specific properties of an individual's speech variability and sociolinguistic identity, such as gender, age, race, class, speaking style, or group membership (e.g., normal talker, or talker with a speech-language disorders). The results of this study indicated that a number of indexical attributes were affected by F0 manipulation, including personal identity (gender and age), clinical variables (cognitive and speech and language abilities), and personal presentation variables (anxiety and likeability).

Age was one such indexical attribute that was affected by F0 modifications that can be said to influence listener's perception of a key identity feature of the speaker. Often, F0 is correlated to the indexical attribute of age, so that it can be expected that younger children have a more naturally raised F0, and older children have a more naturally lower F0. If atypical F0 ranges are observed, it can be said that listeners will naturally judge and perceive the speaker more harshly. The findings reported here reflect this notion in that the Original modification and the Mild severity level reflect the most positive and appropriate ratings based upon the CWSs' chronological age. The trends observed in this study were expected, being that the Lowered modification resulted in perceived older children, and the Raised modification resulted in perceived younger children. This trend is also observed throughout the severity groups, where CWS in the Mild and Severe severity groups were perceived as older than CWS in the Moderate severity group, implying that CWS in the Moderate severity group were perceived as closer to the child's chronological age. Therefore, it can be stated again that the further F0 is from a typical range expected for the speaker, regardless of severity level, the less positive listener perceptions will be associated to that CWS.

Gender was another indexical attribute that was affected by F0 modifications that can be considered a key personal identity feature of a speaker. Similar to age related findings, F0 is oftencorrelated to the indexical attribute of gender and can even affect perceptions of speaker likeability (Perry, Ohde, Ashmead, 2001; Re, O'Connor, Bennett, Feinberg, 2012). The findings reported here showed significant effects of F0 modification on gender judgments; however, F0 modification interacted with severity and will be discussed more below.

Additional indexical attributes that were affected by F0 modification included cognitive abilities and speech-language abilities of CWS. Results indicated that the Original condition led to more positive impressions of CWSs' cognitive abilities (i.e., judgments of less likelihood of impairment) than both the Lowered and the Raised modification conditions. Similarly, the Original modification level was associated with more positive impressions of CWSs' speech-language abilities (i.e., judgments of less likelihood of speech-language impairment) than both the Lowered and the Raised modification conditions. This is consistent with the hypothesis that having an age-appropriate F0 level results in more positive perceptual evaluations of cognitive and speech-language abilities.

Further indexical attributes regarding presentation variables that were affected by F0 modification included anxiety and likeability. Our results demonstrated that F0 only affected judgments of anxiety for CWS at the Moderate severity level, with children in the Original modification condition being perceived as significantly less anxious than the Raised modification condition. Impressionistic judgments of likeability indicated that overall, CWS in the Original modification were perceived as significantly more likeable than both the Lowered and the Raised modification conditions. These results show that modifying F0 played a significant role on impressionistic judgements of speaker attributes of cognitive abilities, and speech and language abilities. The fact that the impressionistic judgements trended towards more positive for the Original modification conditions is consistent with the hypothesis that having an age-appropriate F0 is important for optimal evaluations of a speaker's likeability and degree of perceived anxiety.

Effects of severity level and interactions with severity

Another main finding of this research was that stuttering severity levels also were associated with significant differences in listener perceptions for all measures examined. The segmental and suprasegmental attributes for which significant differences of severity were demonstrated included fluency, intelligibility, and speech rate. Listeners gave impressionistic ratings of degree of perceived fluency that mirrored the diagnostic evaluations that had been provided by speech language pathologists (mild, moderate, or severe). Thus, CWS in the Mild severity group were perceived overall as being more fluent than CWS in the Moderate and Severe groups, and CWS in the Moderate severity group were perceived overall as more fluent than CWS in the Severe group. The severity levels also affected perceived intelligibility, with CWS in the Mild severity group being perceived as more intelligible than those in the Moderate group, and CSD in the Moderate group in turn being perceived as more intelligible than those in the Severe group. Another suprasegmental variable that was affected by severity was speech rate results indicated that CWS in both the Mild and Moderate severity groups were perceived as speaking faster than CWS in the Severe group. Not surprisingly, these results indicated that the less severe the stuttering severity, the more positive the impressionistic judgements. The findings that perceptual evaluations mirror diagnostic categories provided by speech-language pathologists further serves to validate the methodology used in this study.

F0 modification and severity interacted to influence gender judgments. For mild and moderate severity levels, a Raised F0 yielded more “girl” judgments than an Original F0 level. Significant effects of severity, and interactions with severity, must be taken with a grain of salt, because the distribution of males and females was not controlled across severity levels. In particular, 4/5 (80%) in the Severe group were male, a fact which would be expected to cause speakers in the Severe group to be judged as sounding, on average, more like boys than speakers in the other two severity groups. Thus, effects of severity level probably mask inequalities in the distribution of the gender across severity levels.

Clinically related indexical variables regarding cognitive and speech and language abilities of CWS were affected by stuttering severity levels as well. CWS in the Mild severity group were perceived as having better cognitive abilities than CWS in the Moderate group, who in turn were perceived as having better cognitive abilities than CWS in the Severe groups.

Similarly, CWS in the Mild severity condition were perceived as having better speech and language abilities than CWS in the Moderate group, who in turn were perceived as having better speech and language abilities than CWS in the Severe group. These results again indicate that the lower the severity, the more positive impressionistic judgments and further serve to validate the methodology used in the present study. Although CWS in the lower severity groups overall were perceived more positively, this study further demonstrated that modifications of F0 affected ratings for all CWS, regardless of severity level. Thus, F0 modifications did affect perceptual ratings from listeners overall. Therefore, CWS who are more constrained by the impact of his/her fluency, such as those in the more severe severity groups, could work on altering volitional F0 in order to increase perceptual ratings of speech and self.

Personal presentation indexical attributes were also affected by stuttering severity levels, including impressionistic judgments of anxiety and likeability of CWS. Specifically, we found that CWS in the Moderate severity group only were perceived as having lower anxiety in the Original modification condition than in the Raised or Lowered modification conditions. However, for likeability ratings, CWS were perceived as less likable if they were in the Severe severity group rather than the Mild or Moderate groups. Overall, the results of severity levels suggests that the lower the stuttering severity, the more positive the evaluations in impressionistic judgments of a variety of attributes.

Theoretical implications

These results have significant theoretical implications regarding the hypothesis of interdependencies between various aspects of speech. In particular, the findings support the hypothesis that ratings of suprasegmental, segmental, and indexical attributes are strongly interdependent on each other, rather than being characteristics that can be isolated perceptually from the speech of a talker. This is reflected in patterns of perceptual judgments by listeners. In particular, F0, which is often described as a suprasegmental attribute, had effects on attributes that are variously considered segmental, suprasegmental, and/or indexical. The present findings are the first to reveal this pattern of interdependence among segmental, suprasegmental, and indexical components in perception of speech of a clinical population, namely, CWS. These findings suggest that perceptual bias in judgements of suprasegmental, segmental, and indexical speaker characteristics may generally be automatic. It appears that listeners unintentionally and unconsciously perceive and judge speakers based on a variety of characteristics, with little to no control over such bias (cf. Dilley et al., 2013).

Clinical and practical implications

These findings have significant implications for developmental prosody research, as well as for clinical practice for targeted clinical populations such as CWS. Impressionistic judgements are common in both clinical and real world evaluation of communication, providing for qualitative analysis of communication and social interactions at a more functional level. The present research showed that modifying the F0 of CWS significantly affected the impressionistic judgments and perceptions of listeners. Recall that the spectral manipulation method used here allowed for control of a variety of other acoustic variables across conditions, including duration, speech rate, overall F0 contour, and pausing, as well as factors such as linguistic content and voice quality. Therefore, it can be stated that F0 manipulations alone, significantly influence perceptual ratings for a clinical population of CWS. Our initial hypothesis anticipated a more non-monotonic trend, where increases in F0 would yield consistently better (or worse) impressionistic judgments. However, a number of dependent variables showed non monotonic trends, with more positive evaluations for the Original modification condition on a number of impressionistic dimensions. This may be due to a tendency of subjects to make evaluations partly based on the appropriateness of F0 for a child's judged chronological age. The Original modification condition had age-appropriate F0 levels, while the Lowered and Raised levels were associated with having lower than or greater than age-appropriate F0 levels, respectively. If this hypothesis is correct, then children who naturally have an F0 level which is unusually high or unusually low for his/her chronological age would be expected to be judged more negatively on a variety of dimensions than a child with a typical F0 level for his/her chronological age. This hypothesis further raises the possibility that therapies could be developed which were targeted at altering the F0 level of children, or perhaps adults, when it was atypical. This could involve lowering F0 which was unusually high, or raising F0 which was unusually low. Additional work is needed to test the hypothesis that the typicality of F0 levels relative to chronological age influences impressionistic judgments, and to assess the clinical benefits of therapeutic strategies aimed at F0 modification for enhancing impressionistic judgments of talkers.

Given the generally small effects of F0 on dependent variables in the present study, it seems possible that the greatest clinical significance may obtain for CWS with borderline severities. Some CWS are often overlooked or wait-listed for speech therapy due to being on the border between two diagnostic categories (e.g., normal vs. mild stuttering). For these children, there may be particularly great clinical significance of F0 level, in that changing F0 level might yield different classifications of e.g., the degree of fluency for that child in a way that alters their clinical classification.

In summary, the results of this experiment reveal that naïve listeners are influenced by changes in F0, with the data generally consistent with the hypothesis that listeners have more positive judgments when the F0 level of CWS is age and gender appropriate. Therefore, it can be stated that F0 manipulations alone significantly influence perceptual ratings of clinical populations, and that these effects obtain not only for typically developing children and adults (Dilley et al., 2013) but also for a clinical population of CWS. More research is warranted in this area to determine the clinical significance of F0 modifications in other populations. These findings also support the hypothesis that ratings of suprasegmental, segmental, and indexical attributes are interdependent of each and depend not only on isolated characteristics of a talker's speech, but also on the listener's overall assessment of an individual.

Directions of Future Research

This study has implications for future research and findings in the field of speech language pathology. Both theoretical and clinical implications were observed regarding the interdependencies of suprasegmental, segmental, and indexical characteristics, as well as in the influence of F0 specifically in speech and listener perceptions. This study would be worth investigating more using the clinical population of adults who stutter. The results may reveal more significant differences, as adult ratings from listeners would be more concrete. Often, many participants stated difficulty honestly rating the CWS, primarily with the likeability attribute, due to the young age of the children, it was difficult to assess negatively. It would also be worth investigating further whether aligning F0 values to more age appropriate values specifically would significantly affect impressionistic judgments. More research would need to be done regarding natural F0 for children. Additionally, it would be interesting to utilize video, rather than just speech audio for listener perceptions in order to analyze how secondary characteristics of fluency disorders affect listener perceptions as well. This study could be designed so that participants first hear only audio and provide ratings, and then immediately watch the same speech fragment over video and provide ratings. This may mean that rating scales would have to be tweaked slightly for the video ratings as fluency disorders will be much more observable. Furthermore, parsing apart gender perceptions regarding age appropriate F0, and perceived gender roles in modified speech with more balanced ratios of male: female speakers would also be interesting to explore. An additional area to explore in order to assess the generalization of the results from this thesis, would be to consider altering volitional F0 in natural speaking contexts, rather than in spectrally modified contexts. This would require that CWS be trained to adequately alter F0 within natural ranges for typical speaking situations. Listeners would then be asked to rate CWS on the same nine attributes to provide perceptual ratings on segmental, suprasegmental, and indexical attributes of the speaker. This kind study would be interesting in order to determine whether the results of this thesis study generalize when all other aspects are not held constant such as in natural speaking situations, as compared to spectral modifications where all other factors are uninfluenced. Additional research and design would be necessary. The effects of F0 on perceptions of age would also be a variable for further in-depth investigation in order to assess whether modifications in F0 would significantly alter the perceived age of speakers regardless of segmental content of the speech signal. Overall, the present results demonstrate the there is significant potential to develop clinical interventions based on F0 in order to manage listener impressions related to a variety of segmental, suprasegmental, and indexical properties of speech and talkers. Furthermore, altering the anchors of the rating scales and making the scale smaller, as well as eliminating the extreme attributes associated with the end variables (1 and 8) would be an area of change, in order to eliminate any additional bias that may unintentionally affect listener perceptions.

Table 1: Stimulus Lists

LIST ONE	LIST TWO	LIST THREE
SF_1_O	SF_2_L	SF_3_R
SM_1_R	SM_2_O	SM_3_L
SM__L	SM_2_R	SM_3_O
SM_1_O	SM_2_L	SM_3_R
SF__R	SF_2_O	SF_3_L
SM_1_L	SM_2_R	SM_3_O
SM_1_O	SM_2_L	SM_3_R
SM_1_R	SM_2_O	SM_3_L
SF_1_L	SF_2_R	SF_3_O
SM_1_O	SM_2_L	SM_3_R
SF_1_R	SF_2_O	SF_3_L
SM_1_L	SM_2_R	SM_3_O
SF_1_O	SF_2_L	SF_3_R
SM_1_R	SM_2_O	SM_3_L
SF_1_L	SF_2_R	SF_3_O
SF_2_R	SF_3_O	SF_1_L
SM_2_L	SM_3_R	SM_1_O
SM_2_O	SM_3_L	SM_1_R
SM_2_R	SM_3_O	SM_1_L
SF_2_L	SF_3_R	SF_1_O
SM_2_O	SM_3_L	SM_1_R
SM_2_R	SM_3_O	SM_1_L
SM_2_L	SM_3_R	SM_1_O
SF_2_O	SF_3_L	SF_1_R
SM_2_R	SM_3_O	SM_1_L
SF_2_L	SF_3_R	SF_1_O
SM_2_O	SM_3_L	SM_1_R
SF_2_R	SF_3_O	SF_1_L
SM_2_L	SM_3_R	SM_1_O
SF_2_O	SF_3_L	SF_1_R
SF_3_L	SF_1_R	SF_2_O
SM_3_O	SM_1_L	SM_2_R
SM_3_R	SM_1_O	SM_2_L
SM_3_L	SM_1_R	SM_2_O
SF_3_O	SF_1_L	SF_2_R
SM_3_R	SM_1_O	SM_2_L
SM_3_L	SM_1_R	SM_2_O
SM_3_O	SM_1_L	SM_2_R
SF_3_R	SF_1_O	SF_2_L

Table 1 (cont'd)

SM_3_L	SM_1_R	SM_2_O
SF_3_O	SF_1_L	SF_2_R
SM_3_R	SM_1_O	SM_2_L
SF_3_L	SF_1_R	SF_2_O
SM_3_O	SM_1_L	SM_2_R
SF_3_R	SF_1_O	SF_2_L

This table shows the ordering of the three stimulus lists, created using a Latin-Squares experimental design to counterbalance across all measures. Each stimulus list is divided into 3 blocks of 15 samples each (15 CWS).

Table 2: Attribute Averages for Modification and Severity Conditions

Attribute and Severity	Lowered	Original	Raised
Age-Mild (in years)	6.72	6.99	6.23
Age-Moderate (in years)	5.04	4.27	4.22
Age-Severe (in years)	5.21	4.53	4.77
Anxiety-Mild	3.92	3.75	4.10
Anxiety-Moderate	3.89	3.56	4.15
Anxiety-Severe	4.67	4.42	4.37
Cognitive Abilities-Mild	3.39	2.79	3.16
Cognitive Abilities-Moderate	3.73	3.10	3.22
Cognitive Abilities-Severe	4.42	4.13	4.21
Fluency-Mild	4.93	4.64	5.06
Fluency-Moderate	5.13	4.95	5.36
Fluency-Severe	5.98	5.92	5.97
Gender-Mild (larger = more female)	4.69	4.34	5.41
Gender-Moderate (larger = more female)	4.48	4.30	4.90
Gender-Severe (larger = more female)	3.55	3.20	3.37
Intelligibility-Mild (percentage)	60.99	65.65	56.30
Intelligibility-Moderate (percentage)	53.29	62.87	54.57
Intelligibility-Severe (percentage)	50.18	58.50	49.72
Likeability-Mild	3.95	3.50	3.82
Likeability-Moderate	3.88	3.50	3.75
Likeability-Severe	4.36	4.19	4.19
Speech Rate-Mild	4.56	4.41	4.30
Speech Rate-Moderate	4.63	4.50	4.38
Speech Rate-Severe	5.33	4.38	5.03
Speech and Language Abilities-Mild	4.54	3.99	4.59
Speech and Language Abilities-Moderate	4.85	4.58	4.64
Speech and Language Abilities-Severe	6.27	6.19	5.89

Table 2 shows the average ratings for each attribute examined based upon severity level and modification condition. Recall that higher ratings represent more negative impressionistic judgments, with the exception of age (represented in years and months), intelligibility (represented as a percentage), and gender (higher ratings represent more female-like).

APPENDICES

Appendix A: Participant Instructions and Reference Sheet

In this study, you will listen to 50 short speech fragments over headphones and then answer a few questions immediately following each speech fragment. The questions that follow will ask you to rate the speaker on 9 aspects of the speaker based on an eight point scale. There will be one practice block of 5 speech fragments, followed by a chance to ask questions before the rest of the experiment continues. Please use that time to clarify any questions that you may have.

The first rating that you will provide will be in regards to the age of the speaker. You should estimate the speaker's age in years and months using the format Age.Month (age separated from month by a period). For example: 1 year and 3 months, depicted as 1.3. Recall that children typically begin speaking first words at approximately one year of age, and begin speaking in phrases at approximately 3 years of age.

The second rating that you will provide will be in regards to the gender of the speaker. You should assess your confidence in regards to the gender of the speaker, with 1 indicating that you are positive the speaker is a male, and 8 indicating that you are positive the speaker is a female.

The third rating that you will provide will be in regards to the speech rate of the speaker. You should assess your perception of how fast the speaker is speaking, with 1 indicating that the speaker is speaking fast, and 8 indicating that the speaker is speaking slowly.

The fourth rating that you will provide will be in regards to the intelligibility of the speaker. You should assess your perception of what percentage of words were understood, with

0 indicating 0%, or none of the words were able to be understood, and 100 indicating that 100%, or all of the words were able to be understood.

The fifth rating you will provide will be in regards to the speaker's cognitive abilities. You should assess your confidence in regards to the overall cognitive, or mental abilities, of the speaker, with 1 indicating that you are positive that the speaker has no cognitive impairment, and 8 indicating that you are positive the speaker does have a cognitive impairment.

The sixth rating you will provide will be in regards to the speaker's speech and language abilities. You should assess your confidence in regards to the overall speech and language abilities of the speaker, with 1 indicating that you are positive that the speaker has no speech-language delay or disorder, and 8 indicating that you are positive that the speaker does have a speech-language delay or disorder.

The seventh rating you will provide will be in regards to the speaker's anxiety. You should assess your confidence in regards to the overall anxiety of the speaker, with 1 indicating that the speaker is completely non-anxious/comfortable, and 8 indicating that the speaker is completely anxious/uncomfortable.

The eighth rating you will provide will be in regards to the speaker's likeability. You should assess your confidence in regards to the overall likeability of the speaker, or how enjoyable he/she would be as a conversational partner, with 1 indicating that the speaker is strongly likeable, and 8 indicating that the speaker is strongly unlikeable.

The ninth, and final rating you will provide will be in regards to the speaker's fluency. You should assess the overall degree of fluency of the speaker when speaking, with 1 indicating that the speaker is completely fluent, and 8 indicating that the speaker is completely disfluent.

1. **Age:** Format Age.Month (age separated from month by a period). Example: 1 year and 3 months, depicted as 1.3.
2. **Gender:** Assess the gender of the speaker as male or female, with 1 indicating positively male, and 8 indicating positively female.
3. **Speech rate:** Assess the rate of speech, or how fast the speaker is speaking, with 1 indicating positively fast, and 8 indicating positively slow.
4. **Intelligibility:** Assess what percentage of the phrase was able to be understood, with 0 indicating that 0%, or none of the words were able to be understood and 100 indicating that 100%, or all of the words were able to be understood.
5. **Cognitive abilities:** Assess the overall cognitive, or mental abilities of the speaker, with 1 indicating that you are positive the speaker has NO cognitive impairment, and 8 indicating that you are positive the speaker HAS a cognitive impairment.
6. **Speech and Language abilities:** Assess the overall speech and language abilities of the speaker, with 1 indicating that you are positive the speaker has NO speech and/or language delay/disorder, and 8 indicating that you are positive the speaker HAS a speech and/or language delay/disorder.
7. **Anxiety:** Assess the overall anxiety/comfort of the speaker, with 1 indicating that the speaker is positively non-anxious and comfortable, and 8 indicating that the speaker is positively anxious and uncomfortable.
8. **Likeability:** Assess the overall likeability of the speaker, or how enjoyable he/she would be as a conversational partner, with 1 indicating that the speaker is strongly likeable (would go out of your way to speak to him/her), and 8 indicating that the speaker is strongly dislikable (would not choose to intentionally speak to him/her)
9. **Degree of Fluency:** Assess the overall degree of fluency of the speaker when speaking, with 1 indicating that the speaker is completely fluent, and 8 indicating that the speaker is completely disfluent.

Appendix B: Participant Post-Experiment Questionnaire

1. What is your age?
2. What is your gender?
3. What is your major/area of study?
4. How many years of college have you completed?
5. Describe how much and what kind of interaction you have with children on a regular basis?
6. Describe how much and what kind of interaction you have with children with disabilities (cognitive, speech-language, developmental, or otherwise).
7. Did you notice any patterns throughout the experiment? If so, please describe the types of patterns you may have noticed, and how you figured them out.
8. What were your general impressions regarding the sound quality of the speech files?
9. Did you notice any oddities within the speech samples themselves throughout the experiment? If so, please describe the types of oddities that you may have noticed.

10. Do you think the speech played over the headphones was modified? If yes, please describe.

11. What do you think was the purpose of this study?

Please leave additional feedback and comments on the back of this paper. Thank you for your time.

REFERENCES

REFERENCES

Adams, M. R. (1975). A physiologic and aerodynamic interpretation of fluent and stuttered speech. *Journal of Fluency Disorders, 1,* 35-47.

Anderson, J. D., Wagovich, S. A., & Hall, N. E. (2006). Nonword repetition skills in young children who do and do not stutter. Journal of Fluency Disorders, 31(3), 177-199. ASHA (2014). Stuttering. American Speech Language Hearing Association. Retrieved from HYPERLINK "http://www.asha.org/public/speech/disorders/stuttering/" www.asha.org/public/speech/disorders/stuttering/ on February 2, 2014.

Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of Phonetics*, *26*(1), 3-25.

Beilby, J. M., Byrnes, M. L., & Young, K. N. (2012). The experiences of living with a sibling who stutters: a preliminary study. *Journal of Fluency Disorders, 37*(2), 135-148.

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

Bloodstein, 0. (1995). A handbook on stuttering (5th ed.). San Diego, CA: Singular Publishing Group, Inc.

Bricker-Katz, G., Lincoln, M., & McCabe, P. (2010). Older people who stutter: barriers to communication and perceptions of treatment needs. *International Journal of Language & Communication Disorders, 45*(1), 15-30.

Bruce, G. (1977) Swedish word accents in sentence perspective. Lund: Gleerup

Craig, A., Blumgart, E., Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of Fluency Disorders, 34*, 61-71.

Craig, A., Tran, Y., Wijesuriya, N., & Boord, P. (2006). A controlled investigation into the psychological determinants of fatigue. *Biological Psychology, 72*, 78–87.

Craig, A., & Tran, Y. (2006). Chronic and social anxiety in people who stutter. Advances in Psychiatric Treatment, 12, 63–68.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 14*(1), 113.

Darwin, C. J. (1975). The dynamic use of prosody in speech perception. In A. Cohen & S. Nooteboom (Eds.), Structure and process in speech perception New York: Springer-Verlag.

Dilley, L. C., Wieland, E. A., Gamache, J. L., McAuley, J. D., & Redford, M. A. (2013). Age-related changes to spectral voice characteristics affect judgments of prosodic, segmental, and talker attributes for child and adult speech. *Journal of Speech, Language, and Hearing Research, 56*(1), 159-177.

Ferrer, L., Scheffer, N., & Shriberg, E. (2010, March). A comparison of approaches for modeling prosodic features in speaker recognition. *In Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on* (pp. 4414-4417). IEEE.

Franken, M. C., Boves, L., Peters, H. F., & Webster, R. L. (1992). Perceptual evaluation of the speech before and after fluency shaping stuttering therapy. *Journal of Fluency Disorders*, *17*(4), 223-241.

Garner, W. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum

Gabel, R. M., Blood, G. W., Tellis, G. M., & Althouse, M. T. (2004). Measuring role entrapment of people who stutter. *Journal of Fluency Disorders, 29*(1), 27-49.

Gerstman, L. H. (1968). Classification of self-normalized vowels. IEEE *Transactions on Audio and Electroacoustics, AU-16,* 78—80.

Goldinger, S. D. (1996). Words and voices: episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1166.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. Psychological Review, 105(2), 251.

Goldinger, S. D., & Azuma, T. (2004). Episodic memory reflected in printed word naming. Psychonomic Bulletin & Review, 11(4), 716-722.

Goldinger, S. D., Pisoni, D. B., & Logan, J. S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(1), 152.

Grossman, R. B., Bemis, R. H., Skwerer, D. P., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high functioning autism. *Journal of Speech Language Pathology, and Hearing Research, 53*(30), 778-793.

Halle, M. (1985). Speculation about the representations of words in memory. In V. A. Fromkin (ed.), Phonetic Linguistics (pp. 101-104). New York, NY: Academic Press.

Healey, E. C. (1982). Speaking fundamental frequency characteristics of stutterers and nonstutterers. *Journal of Communication Disorders, 15*(1), 21-29.

Henry, M. J., McAuley, J. D., & Zaleha, M. (2009). Evaluation of an imputed pitch velocity model of the auditory tau effect. *Attention, Perception, & Psychophysics, 71*(6), 1399 -1413.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26*(5), 1570.

Hughes, S., Gabel, R., Irani, F., Schlagheck, A. (2010). University students' explanations for their descriptions of people who stutter: An exploratory mixed model study. *Journal of Fluency Disorders, 35* (2010), 280-298.

Ingham, R.J. (1975). Operant methodology in, stuttering therapy. Eisenson (Ed.), Stuttering: A second symposium. New York: Harper & Row.

Ingham, R.J., & Lewis, J. I. (1978). Behavior therapy and stuttering: And the story grows. Human Communication, 1978, 3, 125-152.

Ingham, R. J., Martin, R. R., Haroldson, S. K., Onslow, M., & Leney, M. (1985). Modification of listener-judged naturalness in the speech of stutterers. *Journal of Speech, Language, and Hearing Research, 28*(4), 495-504.

Ingham, R. J., Montgomery, J., Ulliana, J. (1983). The effect of manipulating phonation duration on stuttering. *Journal of Speech and Hearing Research, 26*, 579-587.

Ingham, R. J., Bothe, A. K., Wang, Y., Purkhiser, K., & New, A. (2012). Phonation interval modification and speech performance quality during fluency-inducing conditions by adults who stutter. *Journal of Communication Disorders,45*(3), 198-211.

Johnson, W. (1959). The onset of stuttering: research findings and implications. Minneapolis, MN University of Minnesota Press.

Joos, M. (1948). Acoustic phonetics. *Language, 24*(2), 5-136.

Kent, R. D. (1983). The segmental organization of speech. The production of speech (pp. 57-89). Springer New York.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America, 87*(2), 820-857.

Klassen, T. R. (2001). Perceptions of people who stutter: Re-assessing the negative stereotype. *Perceptual and Motor Skills, 92*(2), 551-559.

Kleinow, J., & Smith, A. (2000). Influences of length and syntactic complexity on the speech motor stability of the fluent speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 43*, 548–559.

Kuhl, P. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics, 50,* 93-107.

Ladd, D. R., Faulkner, D., Faulkner, H., & Schepman, A. (1999). Constant" segmental anchoring" of F0 movements under changes in speech rate. *The Journal of the Acoustical Society of America*, *106*(3 Pt 1), 1543-1554

Ladd, D. R., Silverman, K. E., Tolkmitt, F., Bergmann, G., & Scherer, K. R. (1985). Evidence for the independent function of intonation contour type, voice quality, and F0 range in signaling speaker affect. *The Journal of the Acoustical Society of America*, *78*(2), 435 444.

McClelland, J. L., & Rumelhart, D. (1985). Distributed memory and the representation of general and specific information. Journal of Experimental Psychology: General, 114, 159-188.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76,* 165—178.

Melara, R. D., & Marks, L. E. (1990). Interaction among auditory dimensions: Timbre, pitch, and loudness. *Perception & Psychophysics, 48(*2), 169-178.

Mullennix, J. W., & Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics, 47*(4), 379-390.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *The Journal of the Acoustical Society of America*, *85*(1), 365 378.

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

Pallier, C. (1996). The role of suprasegmentals in speech perception and acquisition. *Phonological structure and language processing: Cross-linguistic studies*, *12*, 145.

Perkins, W. H. (1981). Implications of scientific research for treatment of stuttering--A lecture. *Journal of Fluency Disorders*, *6*, 155--162.

Perkins, W. H., Rudas, J., Johnson, L., & Bell, J. (1976). Stuttering: Discoordination of phonation with articulation and respiration*. Journal of Speech and Hearing Research, 19,* 509-522.

Perry, T. L., Ohde, R. N., & Ashmead, D. H. (2001). The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America*, *109*(6), 2988 2998.

Pierrehumbert, J. (2003). Phonetic diversity, statistical learning and acquisition of phonology. *Language and Speech, 46*(2-3), 115-154.

Pisoni, D. B. (1993). Long-term memory in speech perception: Some new findings on talker variability, speaking rate and perceptual learning. *Speech Communication, 13*(1), 109-125.

Porfert, A. R., & Rosenfield, D. B. (1978). Prevalence of stuttering. *Journal of Neurology, Neurosurgery & Psychiatry, 41*(10), 954-956.

Prosek, R. A., & Runyan, C. M. (1982). Temporal characteristics related to the discrimination of stutterers' and nonstutterers' speech samples. *Journal of Speech and Hearing Research, 25*(1), 29.

Re, D. E., O'Connor, J. J., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for very low and very high voice pitch in humans. *PloS one*, *7*(3), e32719.

Remez, R. E., Fellowes, J. M., & Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance, 23*(3), 651.

Rickheit, G., Strohner, H., & Vorwerg, C. (2008). The concept of communicative competence. Handbook of communication competence, 15-62.

Riley-Graham, J. (2011). Self-Monitoring and Feedback in Disordered Speech Production.

Runyan, C. M., & Adams, M. R. (1979). Unsophisticated judges' perceptual evaluations of the speech of "successfully treated" stutterers. *Journal of Fluency Disorders, 4*(1), 29-38.

Smith, A., & Kelly, E. (1997). Stuttering: A dynamic, multifactorial model. In R. F. Curlee & G.M. Siegel (Eds.), Nature and treatment of stuttering: New directions (pp. 204–217). Needham Heights, MD: Allyn and Bacon.

Sommers, M. S., Nygaard, L. C., & Pisoni, D. B. (1994). Stimulus variability and spoken word recognition. I. Effects of variability in speaking rate and overall amplitude. *The Journal of the Acoustical Society of America, 96*(3), 1314-1324.

Sommers, M. S., & Barcroft, J. (2006). Stimulus variability and the phonetic relevance hypothesis: Effects of variability in speaking style, fundamental frequency, and speaking rate on spoken word identification. *The Journal of the Acoustical Society of America*, *119*(4), 2406-2416.

Spitzberg, B. H., & Cupach, W. R. (1984). Interpersonal communication competence.

Stager, S. V., Jeffries, K. J., & Braun, A. R. (2004). Common features of fluency-evoking conditions studied in stuttering subjects and controls: an H215O PET study. *Journal of fluency disorders*, *28*(4), 319-336.

St. Louis, K., & Hinzman, A. (1988). A descriptive study of speech, language, and hearing characteristics of school-aged stutterers. *Journal of Fluency Disorders, 13,* 331-355.

Studdert-Kennedy M. Speech perception. In: Lass NJ, editor.Contemporary Issues in Experimental Phonetics. New York: Academic Press; 1976. pp. 243–293.

Tasko, S. M., McClean, M. D., & Runyan, C. R. (2007). Speech motor changes coincident with stuttering treatment. *Journal of Communication Disorders, 40*, 42–65.

Van Riper, C. (1982). The nature of stuttering (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Von Tiling, J. (2011). Listener perceptions of stuttering, prolonged speech, and verbal avoidance behaviors. *Journal of communication disorders, 44*(2), 161-172.

Wingate, M. E. (1969). Sound and pattern in "artificial" fluency. *Journal of Speech, Language, and Hearing Research, 12*(4), 677-686.

Wingate, M. E. (1970). Effect on stuttering of changes in audition. *Journal of Speech and Hearing Research, 13*(4), 861.

Wolk, L., Conture, E. G., & Edwards, M. L. (1990). Coexistence of stuttering and disordered phonology in young children. *South African Journal of Communication Disorders, 37,* 15-20.

Woods, C. L. (1978). Does stigma shape the stutterer? *Journal of Communication Disorders, 11,* 483-487.

Van Riper, C. (1982). The nature of stuttering. Engleulood Cliffs, NJ: Prentice-Hall.

Yaruss, J. S., & Conture, E. G. (1996). Stuttering and phonological disorders in children.Examination of the Covert Repair Hypothesis. *Journal of Speech, Language, and Hearing Research, 39*(2), 349-364.