Listener Ratings and Acoustic Characteristics of Intonation Contours Produced by Children with Cochlear Implants and Children with Normal Hearing

by Ioana Barbu

B.A. In Linguistics & Psychology, June 2012, University of California, Los Angeles

A Thesis submitted to

The Faculty of
The Columbian College of Arts and Sciences
of The George Washington University
in partial fulfillment of the requirements
for the degree of Master of Arts

May 15th, 2016

Thesis directed by

James Mahshie
Professor of Speech and Hearing Sciences

Acknowledgements

This research was made possible through a generous grant from the National Institute on Disability and Rehabilitation Research (NIDRR) H133E80006 and National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR) H133G120272. I wish to extend a special thank you to my thesis advisor, Dr. James Mahshie, for supporting me in the completion of this manuscript. I am very grateful to my thesis committee, Dr. Cynthia Core and Dr. Sangsook Choi, who provided their insight and expertise along the way. A warm thank you to the members of the Cochlear Implant Communication Lab who collected and analyzed data. I am also grateful to Andrei Claudiu Cosma within the School of Engineering and Applied Sciences at the George Washington University for his aid in troubleshooting and further developing the Praat scripts. Finally, thank you to the participants and parents who took the time to partake in this study. Your efforts are appreciated.

Abstract of Thesis

Listener Ratings and Acoustic Characteristics of Intonation Contours Produced by Children with Cochlear Implants and Children with Normal Hearing

Cochlear implants (CIs), although effective in restoring auditory sensation for deaf individuals, are lacking in fundamental frequency (F0 or pitch) and temporal fine structure information. Consequently, many aspects of speech perception are significantly compromised. It is reasonable then to suspect that with limited access to F0 and fine temporal structure of speech, the ability to produce intonation patterns by children with cochlear implants (CWCI) would be affected as well. Therefore, perceptual and acoustic analyses were conducted in order to examine production of intonation patterns by CWCI to signal yes/no question and statement contrasts as compared to an age matched control group of children with normal hearing (CWNH). Fourteen CWCI participated in the study, ranging in age from 3;7 to 7;5 years; and 14 CWNH were between the ages of 3;4 – 7;4 years. Statements and questions were elicited using an innovative methodology during a role-play session and were digitally recorded. The elicited productions were parsed, separate files were created for each utterance, and then utterances were randomly presented to a group of 10 normal hearing adult listeners via headphones. Listeners rated the intonation pattern of each production as ranging from falling to rising using a visual analog scale displayed on a computer screen. These represented the listener judgments data and analysis. For the acoustic analysis, the final two syllables of each utterance were identified and the beginning and end of each vocalic portion of the syllable (VPS) was marked using Praat software Version 5.3.51 (Boersma and Weenink, 2013). Mean F0) and intensity measures of the VPS were extracted. The results from the listener judgments task revealed that CWCI and CWNH could distinctively produce rising and falling intonation contrasts to signal a question or a statement. Results from the acoustic analyses suggested a systematic distinction in F0, and to a lesser extent, in intensity, between statements and questions. Examination of the relation between acoustic characteristics and adult listener perceived judgments of intonation revealed large, significant relationships between listener judgments and final syllable F0 as well as F0 and intensity changes between the final and penultimate syllables. Future research directions and clinical implications for evaluation and intervention of prosody are discussed.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

Literature Review

Spoken language can be described and categorized in terms of segmental and suprasegmental characteristics. Segmental, or structural aspects of speech are often described by the phonemes, syllables, words, phrase units, or sentences within a language (Crystal, 1986). Suprasegmental, or prosodic aspects of speech include pitch, stress, intensity, and quantity information (Lehiste, 1976). These suprasegmental features can range over several segments (Tserdanelis & Wong, 2004), which may indicate that the boundaries of the prosodic aspects such as pitch, intensity, and duration or quantity are not always clear. Nevertheless, past research suggests that discrimination of suprasegmental aspects of spoken language is learned early on in life by children with normal hearing (Loeb & Allen, 1993). It can then be assumed that a disruption to typical development, such as a hearing impairment, may affect this aspect of language.

Since their introduction in the 1980s (ASHA, 2004) cochlear implants (CIs) have had an increasingly positive impact on the speech perception and production of children with severe-to-profound hearing loss. CIs have allowed children with little to no hearing to essentially gain ‘hearing’ through audition, which has in turn helped them learn and produce intelligible speech (Nittrouer, 2010). However, CIs provide only limited access to fundamental frequency (F0) – an important cue used to signal intonation contrasts in rising and falling pitch between yes-no questions[1] and statements. Torppa et al. (2014) further suggests that CWCI may be at a disadvantage when compared to their age-matched peers in the perception of intensity as well as F0 due to a lack of fine-grained spectral and temporal detail in the CIs. Given limited access to important intonation cues,

---

[1] From this point onward, “yes-no questions” will be referred to as “questions” in the text.

it is unclear how effectively children with cochlear implants (CWCI) are able to learn to produce intonation contrasts. The devices do, however, provide less direct cues about F0 obtained from temporal cues in the fine structure (Geurts & Wouters, 2001; Green, Faulkner, & Rossen, 2004) as well as intensity and duration cues from the speech envelope that have been shown to relay prosodic information (Lehiste, 1976). Recent work examining segmental perception and production further suggested that CWCI are able to learn about speech production through such factors as vision, instruction, or linguistic knowledge without direct access to speech elements through audition (Mahshie, Core, & Larsen, 2015). It is feasible that reliance on alternate cues might enable CWCI to learn to control F0 to produce intonation distinctions. Alternatively, CWCI may produce these contrasts with less reliance on pitch or intensity variability.

Pre-lingually deafened children are of interest specifically because they do not have access to the same intonation input or experience as children with normal hearing (CWNH). CWCI may therefore be at a disadvantage since they do not have access to auditory information early in development, and could be expected to have deficits or delays in prosodic aspects of speech such as intonation development. Most research suggests that the children with cochlear implants do not perceive intonation patterns to the same extent as their normal hearing peers (Peng, Tomblin, & Turner, 2008; See et al., 2013; Deroche et al., 2014). One consistent finding has been that speech perception abilities of CWCIs improve with an increased amount of experience using the device (Tyler, Teagle, Kelsay, Gantz, Woodworth, & Parkinson, 2000; Fryauf-Bertschy, Tyler, Kelsay, Gantz, & Woodworth, 1997). Even when intonation patterns begin to approximate those of hearing children, the more limited access to important prosodic cues, it may result in CWCI's productions differing from those of CWNH. That is, there may be reliance on different cues than those used by CWNH to distinguish between questions and statements. Listeners might, in turn, rely on different cues in perceiving these distinctions between questions and statements. Further investigation is needed of both the acoustic characteristics of question and statement production by CWCI as well as the relationship between these measures and listener judgments.

There are few studies that discuss production of intonation patterns by implanted children. For instance, Allen and Arndorfer (2000), relied on data from children with hearing aids to draw conclusions about the production of hearing impaired children. They found that 4 out of 6 children with hearing aids could produce acoustically different intonation contours for questions and statements. Although the children with hearing aids used acoustic cues such as F0, intensity, and duration to mark the contrast, the "contrastive use of the cues was less pronounced" than that of their normal hearing peers. Similar findings are supported by studies completed by Peng et al. (2007, 2008), in reference to production of questions and statements of implanted children. These studies looked at the productions of questions and statements of pediatric CI recipients and employed an imitation task and an elicited picture description task. Both studies concluded that the CI group was not consistent in raising pitch to the same extent as their normal hearing peers. In contrast to findings from productions of questions and statements in children with hearing aids (Allen & Arndorfer, 2000), more recent studies found that children with cochlear implants outperformed children with hearing aids in productions that addressed phrasing and pitch (Lenden & Flipsen, 2007), perhaps rendering the initial results questionable.

Tone languages such as Vietnamese or Mandarin Chinese have pitch patterns that mark lexical distinctions between syllables (e.g. *ma* might mean ‘mother’ or ‘scold’ depending on whether it is produced with a high-level tone or with a high-falling tone). Recent research looking at lexical tone production and perception of pre-lingually deaf children with CIs found that a majority of them did not master the production of Mandarin tones (e.g., Peng et al., 2004). Specific difficulty was noted with production of rising intonation as might be encountered in a question, versus a falling intonation encountered in a statement. This has been a consistent finding across multiple studies, and it is one that follows intonation development patterns of children with normal hearing (Cruttenden, 1981; Xu & Sun, 2002). Nevertheless, these children were between the ages of 6;0 and 12;6, with variable implantation ages (2;3 to 10;3). Another Mandarin tone recognition study by Wang et al. (2013), looked at adult tone perception in normal hearing, hearing impaired listeners, and post-lingually deafened users with CIs. Although the results indicated accuracy above 80% for each listener, perception of a high-rising tone was most difficult for the CI recipients and their accuracy of tone recognition was the lowest within the populations tested (e.g. 81% versus 99.3% for normal hearing, 96.4% for individuals with moderate hearing impairment, 93.7% for individuals with moderate-to-severe hearing impairment, and 83.9% for individuals with severe hearing impairment). These findings are similar to those by Peng et al. (2004), in that Tone 2 (high-rising) is most challenging to distinguish in perception (and in production).

Some limitations among the research discussed and cited above, however, include small samples, with a broad range of ages and duration of experience with their implants. Recent findings suggest that early implantation may provide a significant advantage for pre-lingually deafened children in areas of speech perception. There are a number of studies that have suggested that age of implantation is an important determiner of successful outcomes (Connor, 2006; Geers, 2002; Miyamoto, 1994) with earlier implantation resulting in improved speech and language performance (Zhu et al., 2011; Connor et al., 2006). Research by Zhu et al. (2011) and Connor et al. (2006) indicate then that children with earlier age of implantation may provide new information regarding speech perception and production. However, no known studies were conducted on CWCI who have had early implantation in regards to their production of rising and falling pitch patterns, more specifically in natural occurring speech. Further research is indeed needed in order to better understand how CWCI with relatively early age at implantation may use the available input to produce intonation contrasts and how this compares to the production patterns of CWNH.

Research Aim, Goals and Hypotheses

The aim of the present work was to describe the characteristics of rising and falling intonation patterns produced by CWCI using listener ratings and acoustic measures compared to those of typically hearing children. Thus the two populations represented were CWCI who had been implanted prior to 3 years of age and age-matched CWNH. The study used a mixed design with a between-subject factor of hearing status (CWCI vs. CWNH), and a within-subject factor of elicitation types (questions vs. statements). The dependent measures included listeners' judgments on rising and falling intonation patterns, and measurements of acoustic attributes such as fundamental frequency (mean F0) and intensity (mean Intensity).

The goals of the proposed project were:

1. To characterize and compare listener ratings of elicited questions and statements produced by CWCI and CWNH. More specifically, this research goal addressed the following questions:

   1a. Do CWCI produce an intonation distinction as perceived by listeners in utterances elicited as questions or statements?

   1b. Do CWNH produce an intonation distinction as perceived by listeners in utterances elicited as questions or statements?

   1c. How do listener judgments of intonation by CWCI compare to CWNH?

2. To characterize and compare the acoustic characteristics of elicited questions and statements produced by CWCI and CWNH. That is, the research addressed these questions:

   2a. To what extent do CWCI produce distinctions between elicited questions and statements as reflected in acoustic measures of frequency and intensity?

   2b. To what extent do CWNH produce distinctions between elicited questions and statements as reflected in acoustic measures of frequency, and intensity?

   2c. How do acoustic measures of frequency and intensity of these utterances by CWCI compare to those of CWNH?

3. To examine the relationship between listener judgments of intonation patterns and acoustic characteristics of the utterances judged. Finally, this research addressed the following questions:

   3a. How do results from listener ratings of intonation patterns relate to acoustic measures of the utterances judged?

Although it has been suggested that numerous variables such as parental involvement, therapeutic intervention, CI wear time, innate cognitive and linguistic abilities, age of implantation, CI technology, early diagnosis of hearing loss, and proper audiologic management may influence outcomes (Connor, 2006; Geers, 2002; Miyamoto, 1994), a majority of the previous studies did not control for these. We hope that a study with participants that have had early implantation and are younger in chronological age may reveal more information about production of intonation contours for children with hearing loss. Given our discussion of previous findings, it is anticipated that CWCI will be able to produce yes-no question and statement contrasts, but not as discernible as CWNH. Given the limited access to F0 information, CWCI would rely more heavily on other types of acoustic cues such as intensity, which is more readily accessible through CIs. Variations within CWCI are expected across implantation age and amount of experience with the device.

Chapter 2: Methods

Participants

Background. Fourteen CWCI (9 females; 5 males) and 14 CWNH (10 females; 9 males) were recruited and data was collected prior to, or concurrent with, this thesis. Participants were part of a larger project examining CWCI's perception and production of various aspects of speech. Methodology and instrumentation were approved by the George Washington University's institutional review board (IRB) and informed consent obtained from the children's parents prior to data collection. Trained research assistants collected the data in one to two sessions, lasting approximately 1-2 hours at the George Washington University Speech and Hearing Clinic, the children's homes, or at the school that the child attended.

Population.

CWCI. Fourteen pre-lingually deafened monolingual English children made up the participant pool for the implanted CWCI. The children participating in this study were between the ages of 3;7 and 7;5, (mean age = 5;0) had an onset of hearing impairment before 1 year of age, demonstrated a profound hearing loss as documented through medical history and an audiogram, used speech as the primary communication modality, had early implantation, and had no additional disabilities. The group of CI children had an average age at implantation of 1;5 years, with 11 CWCI being implanted before age 2, and 3 CI children receiving their implants after turning 2 years old - at 2;2, 2;5, and 2;11 years of age. The Ling 6 Sound Test (Ling 2002) was completed before beginning testing with each CI participant. The test was presented using live voice, and required the children to repeat six different speech sounds from low to high frequency. This ensured that the child had access to the full range of speech sounds necessary for conversational speech. Refer to Tables 1A and 2A in Appendix A for information on the characteristics of the participants.

CWNH. Fourteen monolingual English normal hearing children were recruited through procedures similar to those used to recruit CWCI. The CWNH were age matched to the CWCI, with a mean age of 5;1 and 5;0, respectively. The group of normal hearing children were between 3;4 and 7;4 years of age, presented without a clinical history of speech or language impairment, and had normal hearing as per initial parent report. A hearing screening was completed for each of the CWNH prior to participation in the experiment (20 dB at 1000 Hz, 2000 Hz, 4000 Hz) in order to ensure that there was no hearing loss, and that the child had access to a range of frequencies. Refer to Appendix A for a list of the participants' ages at time of testing.

Hearing Listener Judges. Ten listener judges were recruited from the George Washington University community in order to rate the intonation patterns of CWCI and CWNH. All listener judges were adults with no reported hearing impairments. A hearing screening (25 dB at 1000 Hz, 2000 Hz, 4000 Hz) was completed before the listening experiment began to ensure hearing was within normal limits.

Stimuli

Intonation Production Task. An elicited intonation task was used to obtain speech tokens representing productions of statements and questions by 28 children.

Equipment & Instrumentation. The children's speech was digitally recorded using a Sony 12.1 MP High-Def Camera with 120GB HD and 3.2" LCD Screen. Two different but comparable omni-directional lavalier microphones (WLX-PRO VHF Wireless Lapel Microphone System, and Sennheiser SEEW112PG3A-ew 112-p G3 model) were used. These were attached to a vest, approximately 6 inches from the child's mouth. The recordings were subsequently transferred to a computer using Quicktime Player as MP4 Audio Visual (AV) files. Next, the audio portion of each AV file was exported as a .wav file with a data rate of 128 kbps, as a Stereo channel, sampled at 44.100 kHz, with a bit rate of 16 bits.

Elicitations. To elicit questions or statements, a role-play activity with a hand puppet named Bo was developed (Mahshie, Core, Choi, & Dewey, 2014). This task was refined in order to address limitations of previous methodologies. Previous research based their elicitations on imitation, reading, and/or picture description (Allen & Arndorfer, 2000; Patel & Grigos, 2006; Peng et al., 2007). Although these methods may be appropriate for older children and adults, they were not appropriate for a younger population who either could not read, or would not participate in an imitation/picture description study. The task is ideal for younger children because it relies on more conversational productions with manipulable objects (e.g. foam fruits and vegetables). Unlike previous research methodologies, this task has high ecological validity. Questions and statements are extracted from natural speech as produced by the children during play. It was concluded that the task at hand was more representative of these younger children's speech productions than alternative approaches to eliciting productions.

When completing the task, a clinician used a hand puppet named Bo in order to elicit utterances from all of the children. In order to elicit questions, the subjects were required to, “Ask Bo what he likes to eat.” To elicit statements, the subjects were asked, “What do you think he wants [next]?” If no response or an inappropriate response (e.g. ambiguous) response was provided, the clinician would ask the child to, “Tell Bo what

[he] likes to eat." Each experimenter followed a standard script, which also included modeling for the child if there were no initial responses. At least 4-5 useable[2] questions and 4-5 useable statements were elicited per each child.

The recorded elicitations were parsed using Phon (Yvan et al., 2006), an open-source software used for phonological and phonetic transcriptions and analyses. Phon enables auditing of individual utterances as well as transcription and annotation of utterances in different tiers (a set of rows used to organize and mark segments). Two trained research assistants examined, transcribed, and characterized the elicited utterances. The first transcriber set up and initially entered information into the following five tiers, Tier 1: orthographic transcription; Tier 2: intended elicitation (statement vs. question); Tier 3: used for analysis (included or excluded from analysis); Tier 4: reason for exclusion, when applicable; and Tier 5: miscellaneous notes. The second transcriber checked the information entered by the first transcriber, and noted any disagreements regarding the characterization of the elicited utterances in a separate document. A third trained transcriber acted as the final judge for the coding of the utterance, if a discord took place between the first two transcribers in the tier ratings. Each utterance was then exported to a separate .wav file in preparation for the acoustic analyses. This process ensured that the utterances were useable.

In some cases, utterances were excluded. Productions were excluded if they met any of the following conditions:

- were comprised of only one syllable (e.g. "yea," "no," "that")
- included the child's name
- had competing background noise
- were overlapping with the examiner's voice
- were whispered

---

[2] Refer to exclusion criterion below for determination of "useable" utterances.

- were cut off (e.g. “I would like some ice cr…” where “cream” is clearly missing)
- had atypical voice quality (i.e., using a silly voice)
- were viewed as ambiguous (see below)

Utterances were considered ambiguous if they met the following three criteria, (1) the intonation output did not match the expectation given the context (e.g. the experimenter asked, “What do you think Bo wants to eat next?” and the child responded, “An apple?”); (2) the child demonstrated evidence of producing a question or a statement with the expected intonation patterns as evidenced by other elicited utterances (e.g. the child could produce rising intonation when asking a question); and (3) the child used a behavior that reflected an intention other than that evident in the utterance (e.g. the child held food towards Bo in a gesture reflecting questioning but produced a statement).

Preparation of utterances for analysis. In order to present the utterances for intonation judgments by listeners, one second of silence was added to the beginning of the utterances using Adobe Audition (Adobe, 2014), a sound editing software. The one second of silence was added in order to ensure that listeners had time to respond to an utterance once they advanced through the stimuli during the listening portion of the study. Elicited questions and statements were subsequently processed so that the peak amplitude of each file was normalized to 65 dB RMS amplitude using Level *32* (Carrell, 1998), a batch sound-processing package. This was done in order to ensure that there were no overall amplitude differences among productions that were obtained at different times in potentially different settings (e.g. a classroom versus the clinic rooms).

Procedure

Intonation Analysis.

Listener judgments of Intonation. The prepared utterances were subsequently presented using the presentation software associated with the Soundswell wave editing software package (Hitech, 2008). Preparation of the files involved conversion of the files to .smp (a file format used with Soundswell), and organization into randomized blocks of 40-60 stimuli for a total of 360-400 stimuli. The audio files within a block were prepared for presentation (or “glued”) together using the Glue (Hitech, 2008) program, a software application used specifically for creating blocks of audio files for subsequent presentation and judging. Block presentation for each listener judge was counterbalanced for presentation. In order to determine the consistency or reliability of each listener judge with his or her answers, an additional block was presented. This reliability block was comprised of 10% of the total stimuli, and it was presented within the counterbalanced list of utterances.

Once the blocks were prepared, the audio recordings of each utterance were presented using Judge (Hitech, 2008). The hearing adult listeners were asked to rate the intonation pattern of each utterance as ranging from falling to rising by placing a movable pointer along a visual analog scale displayed on a computer screen (see Figure1). Next, the listener judges were asked to rate their confidence levels regarding their intonation pattern ratings also by moving a pointer along a second visual analog scale. The location of the pointers corresponded to a value, which, depending on the scale ranged from ‘0’ (extremely falling intonation or not confident) to ‘1000’ (extremely rising intonation or very confident). While confidence data was obtained, it was not used in the present study.

The individual utterance ratings from each listener judge were then transferred to an Excel spreadsheet, and reliability scores were calculated. For each listener, ratings within the reliability block were compared to the ratings of the same utterances that were randomly assigned to different blocks. A correlation between these two scores of 80% or better indicated that a listener judge was reliable. The ratings from ten highly reliable listeners (>80%) were averaged and the mean scores were used to characterize the perceived intonation for each utterance. Each listener judge rated 4 to 5 questions, and 4 to 5 statements per child. Refer to Appendix B for a list of the rated utterances.

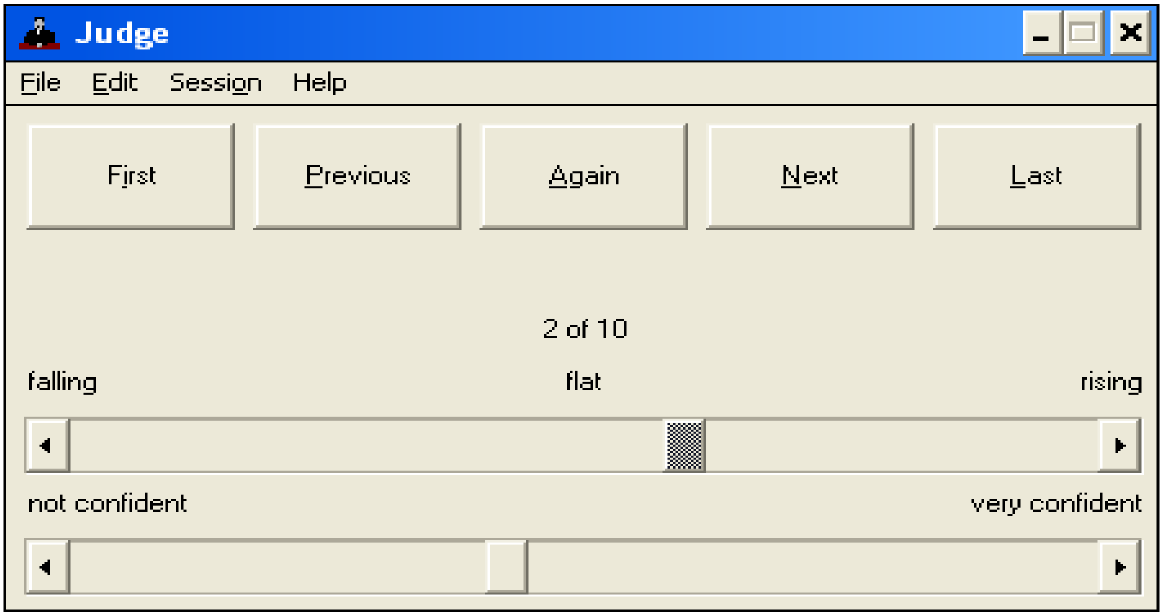

Figure 1. Visual analog scale used to make judgments about utterances using two 0-1000 scales juxtaposing the listeners' perceptions of the children's intonation production patterns with the listeners' confidence levels. Upon completion of the study, reliable listener ratings were recorded.

Acoustic Measurements.

Selection of analysis samples. The recorded utterances used in the listener judgments task were also analyzed acoustically. Based on previous research, the .wav files were down sampled to 22kHz from 44kHz using Adobe Audition (Titze & Titze, 1993; Patel & Grigos, 2006). This was completed in order to produce results that would be more readily comparable to previous studies. A total of 248 utterances were analyzed. Appendix B displays a list of the utterances that were audited and acoustically analyzed. For all acoustic measures, Praat (Boersma & Weenink, 2013) was used. Praat is an open source sound wave editing and analysis software package that provides a set of tools to be used in the analysis of F0 and intensity of a sample.

Intonation conveying portion of the utterance. The initial task was to identify the portions of the utterances that were most likely to convey intonation information to listeners. In previous studies, the vocalic nuclei were considered to be the vowel portion of a syllable and were used in order to gather acoustic measurements such as F0, intensity, and duration, for yes-no questions and statements (Allen & Arndorfer, 2000; Patel & Grigos, 2006; Peng et al., 2007). Some findings indicated that sentence-final F0, the slope between the target word's initial and final F0, duration of the target word, and intensity difference between the penultimate and final syllables were predictive measures for identifying questions and statements in productions by children with hearing impairment (Allen & Arndorfer, 2000). Patel and Grigos (2006) relied on acoustic measures of mean F0, intensity, and duration as well as slope F0 in the final four syllables of the utterances used in order to characterize normal hearing children's contrasts of questions and statements. Peng et al. (2007) reported that final word measures such as mean F0 and mean intensity for the penultimate and final syllables were the most revealing acoustic cues for marking the pitch contours, along with the ratio of the final word to the entire utterance for duration, and slope F0 in the final word of an utterance.

The research studies used the “vocalic nuclei” or the vowel-portion of each syllable in order to conduct measurements of prosody such as F0, intensity, and duration. These studies did not discuss, however, the tendency for intonation to also be conveyed by voiced portions of the syllable that preceded or followed the vowel of a syllable. In connected speech, it was noted that the vowel-only portion of the syllable was often brief and did not capture the prosodic pattern employed by the speaker. To adequately capture the prosodic information of the syllables of interest, it was subsequently decided that the F0 and Intensity measures would be obtained for the entire voiced portion of the PS and FS. That is, the portion of the PS and FS from which F0 and Intensity measures were obtained included vowels and sonorant consonants such as liquids, glides, and nasals that preceded or followed the vowel in each syllable. This portion of the utterance is referred to as the vocalic portion of the syllable (VPS).

Textgrid Annotation. Praat was used to manually identify and label regions to be analyzed for each speech token using a tier system (see Figure 2). We created two tiers – “Orthography,” and “VPS.” The orthography tier included a transcription of the entire utterance. The VPS tier was comprised of two intervals, the first corresponding to the penultimate syllables, and the second corresponding to the final syllable.

The VPS in the penultimate and final syllables were identified based on standard speech segmentation guides (Smith, 2013; Wright & Nichols, 2009; Shriberg & Kent, 1982). When annotating textgrids, we used visual inspection of the intensity tracking capabilities within Praat, auditory perception of the utterance, and spectrographic and waveform indications of periodicity (see Figure 2) to mark the beginning and end of the vocalic portion of the PS and FS.

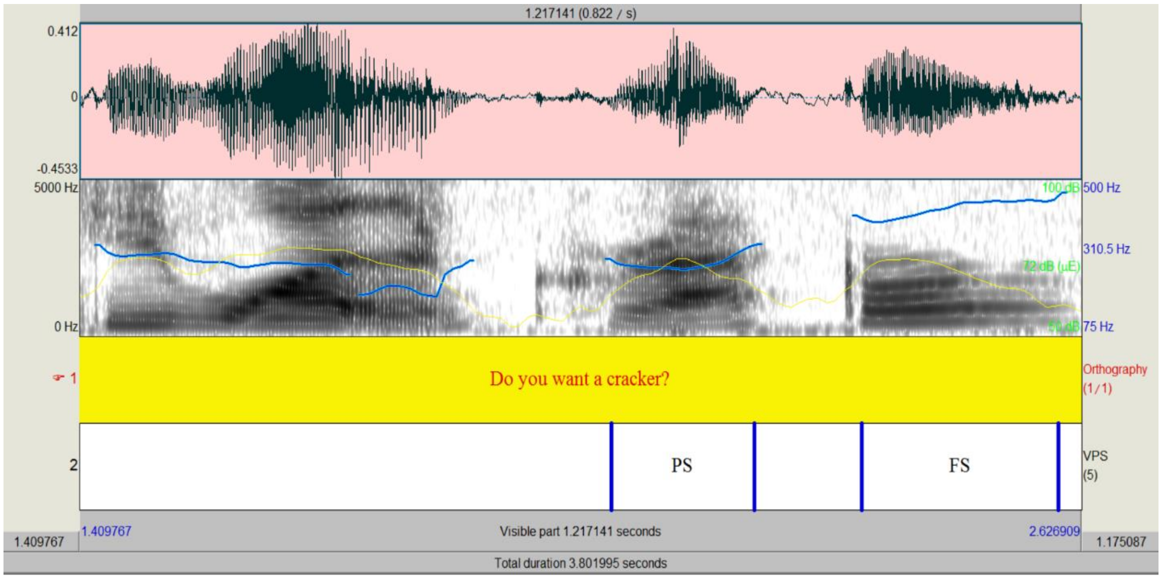

Figure 2. A sample of a segmented utterance: “Do you want a cracker?” The orthography tier includes the transcription of the utterance. The VPS tier contains vowel and vowel-like (e.g. glides) sounds.

Semi-automatic Fundamental Frequency (F0) determination. In order to automate the data analysis process and to increase efficiency, two Praat scripts (a written set of commands meant to help automate measurements) were identified and modified in order to automatically measure the F0 and intensity for each voiced portion (Lennes, 2003; Kawahara, 2010). Each modified script produced a series of pitch and intensity measures. The pitch measures were extracted at 10 millisecond intervals throughout the VPS of each utterance (for PS and FS). Once the scripts were run, Praat generated an output file in a text format. The values from these text files were exported to excel and arranged into columns. Refer to Appendix C for the modified pitch script.

Initially, the measurements were gathered using standard parameter settings. The parameters were changed when the script’s pitch extraction rendered undefined pitch values or invalid measures.

Parameter Optimization & Correction of Undefined Pitch Periods. Two judges sampled 10% of the utterances (1-2 from each speaker) used in the analysis in order to verify the results of the script and to establish the critical parameters for F0 extraction. The following standard procedure was established, and later applied by the primary researcher to the remaining utterances. First, the judges visually inspected undefined or unusual pitch measures on the spectrogram. Next, they listened to each file. Then, they measured the peak-to-peak mean F0 measurements in the acoustic wave, and compared these to the F0 values generated by the script. If needed, the pitch settings in Praat were adjusted to ensure that the program was accurately detecting the pitch. Lastly, the corrections were recorded, and new mean F0 measured were calculated.

The judges also noted optimized settings for challenging data (e.g. those that included productions of creaky voice). It was found that voicing threshold, pitch range, octave cost, and octave jump cost were the most relevant parameters for optimization. The following standard settings were used and possible manipulations noted:

Pitch Range: 100Hz – 600 Hz (manipulations: 40Hz – 1000Hz)
Voicing Threshold: 0.45 (manipulations: 0.15 – 0.75)
Octave Cost: 0.01 (manipulations: 0.01 – 0.1)
Octave Jump-Cost: 0.35 (manipulations: 0.15 – 0.95)

Where feasible, a single set of parameters was used in a script for a given speaker. There were cases, however, where parameter settings needed to be adjusted for each individual syllable (PS, FS). Where parameter adjustments needed to be made for each individual VPS within an utterance, measurements were completed manually following the procedure described above and entered into the spreadsheet for a particular cycle. A total of 70% (173 out of 248) of the utterances required manipulations of at least one out of the four parameters mentioned (pitch range, voicing threshold, octave cost, octave jump-cost) in either PS, FS, or both.

Intensity. Employing an additional modified Praat script, root mean square (RMS) Intensity (dB) values were extracted for each VPS (PS_meanINT, FS_meanINT). In order to verify the intensity script output, two researchers sampled 10% of the utterances and compared the output to manual measurements, completed using Praat's drop down menu for intensity (e.g. "Get Intensity" command). No corrections were required for this output. Refer to Appendix D for the modified intensity script.

Statistical Analysis. Mean F0, mean intensity, and derived differences (F0 difference between PS and FS, Intensity difference between PS and FS) values were obtained for the 248 utterances obtained from 14 CWCI and 14 CWNH. The results were analyzed using SPSS 22 (IBM Corp., 2013). We used a two-way ANOVA in order to explore listener judgments of elicited questions and statements in CWCI and CWNH. The subject group (CWCI vs. CWNH) and elicitation type (Questions vs. Statements) were the two factors. We employed a three-way mixed ANOVA in order to determine between-group (CWCI vs. CWNH) differences and differences between the elicitation types (Questions vs. Statements) on measures of mean F0 and mean Intensity. Finally, the relation among acoustic measures and listener ratings of intonation were examined employing a series of correlations.

Chapter 3: Results

Results are presented in three separate sections related to the three primary research goals posed. First, we present and describe adult listeners' ratings of the elicited questions and statements in implanted children and normal hearing children. Next, we present and describe the acoustic measurements results for the penultimate and final syllables for F0 and intensity. Finally, we explore the relationships between listener judgments and acoustic measurements.

Adult Listeners' Ratings

Mean listener judgments for elicited questions and statements for the CWCI and CWNH are presented in Figure 3. As noted previously, each utterance was rated by 10 judges with judgments resulting in a score ranging from 0 to 1000 (0 = extremely falling, 1000 = extremely rising). The criterion measure for each utterance was the mean score across the ten listeners. The mean rating for statements by the CI children was 403.98 (*SD* = 199.99), and 339.83 (*SD* = 111.97) for the hearing children. For questions, the mean rating was 707.88 (*SD* = 195.82) for implanted children and 739.54 (*SD* = 202.4) for the hearing children.

The data was statistically analyzed using a two-way ANOVA. The independent variables included two factors, group (CWCI vs. CWNH) and elicitation type (question vs. statement). The dependent variable was the mean listener rating scores (0-1000). The two-way ANOVA revealed that there was no significant group difference ($p = 0.88$, see Figure 1). However, there were significant main effects of elicitation type ($F_{(1,\ 56)} = 84.830$, $p < 0.0001$). That is, listeners rated statements with lower rating scores than questions on the visual analog scale. There were no significant interactions between group and elicitation type (*p* = 0.231). These findings suggest that adult normal hearing listeners discriminated between implanted and hearing children's productions of yes-no questions and statements, by rating questions higher than statements.

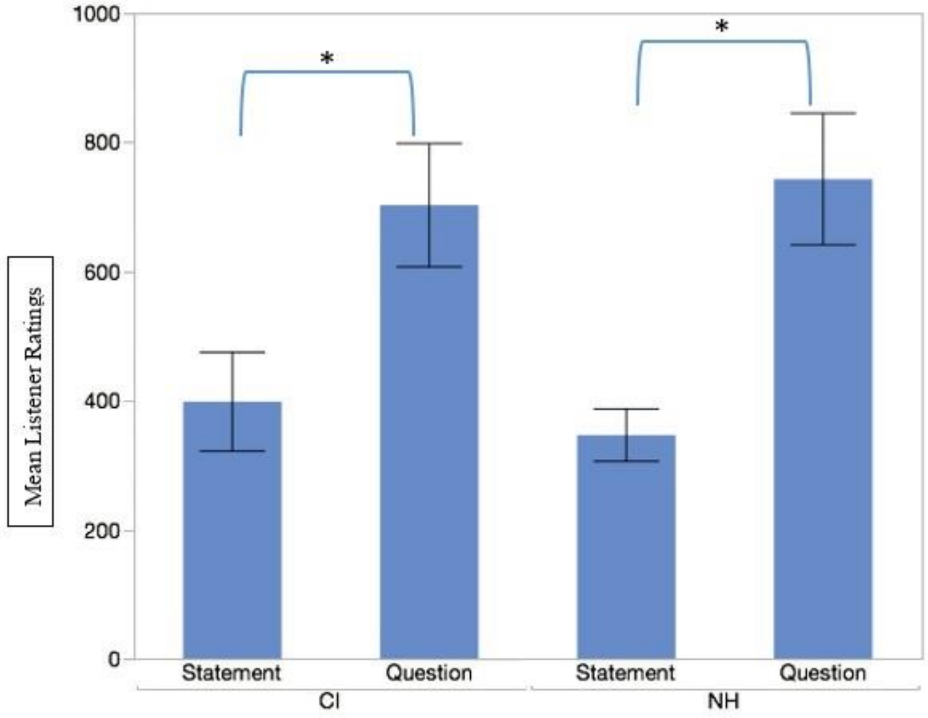

Figure 3. Mean Listener rating scores for elicited statements and questions, in CWCI and CWNH. Error bars represent 95% confidence intervals.
Note. * $p < 0.0001$

Acoustic Measurements. First, given the wide range of ages of the participants in each group (CWNH, 43 – 89 months; CWCI 40-88 months), an inspection of any age-related effects evident in the acoustic measures was necessary. In order to inspect the impact of age on the acoustic measures gathered, a correlation analysis was conducted. Accordingly, age was compared with each of the acoustic measures, as well as the derived measures reflecting a difference in mean intensity between statements and questions and difference in mean f0 between statements and questions. There were no significant correlations between age and any of the acoustic measures. As a result, no adjustments for age were needed in the analysis.

Fundamental Frequency (F0). Figure 4 shows the mean F0 for both the penultimate and final syllables for elicited questions and statements as produced by implanted children and hearing children. The mean F0 for statements by the CI children was 267.86 Hz (*SD* = 73.85) for PS and 241.06 Hz (*SD* = 76.8) for FS. The mean F0 for statements by hearing children was 236 Hz (*SD* = 61.72) for PS and 196.08 Hz (*SD* = 82.05) for FS. For questions, the mean F0 was 263.55 Hz (*SD* = 71.23) for PS, and 299.8 Hz (*SD* = 82.04) for FS, in implanted children. In hearing children, these measurements were 236.06 Hz (*SD* = 66.45) for PS, and 272.65 Hz (*SD* = 89.96) for FS.

A three-way mixed ANOVA was conducted to analyze the F0 measures. The between factors were group (CWCI vs. CWNH) and elicitation type (question vs. statement). The within factor was syllable position (PS vs. FS).

The ANOVA revealed a significant main effect for group, ($F_{(1, 237)} = 14.668$, $p < 0.0005$) and elicitation type, ($F_{(1, 237)} = 14.533$, $p < 0.0005$), but there was no significant main effect for syllable position ($p = 0.153$). There was, however, a significant interaction between syllable position and elicitation type ($F = 54.609$, $p < 0.0001$) while no significant interaction was found between syllable position and group ($p = 0.458$). No significant three-way interactions among elicitation type, syllable position, and group were found ($p = 0.47$).

To explore the interaction between syllable position and elicitation type, a series of post-hoc analyses using t-tests were conducted, and alpha was adjusted using a Bonferroni correction method ($p = 0.0125$). The paired t-test revealed a significant difference in mean F0 between PS and FS, for questions ($t_{(1, 127)} = -5.948$, $p < 0.0001$) and for statements ($t_{(1, 113)} = 4.607$, $p < 0.0001$). That is, for both the CWCI and CWNH, F0 of the FS was less than that of the PS for statements, while F0 of the FS was greater than that of the PS for questions.

The findings suggest that even though CWCI produced higher F0 overall, both groups (CWCI & CWNH) differentiated between questions and statements by producing overall falling pitch changes in the final two syllables for statements and rising pitch changes for questions (see Figure 5).

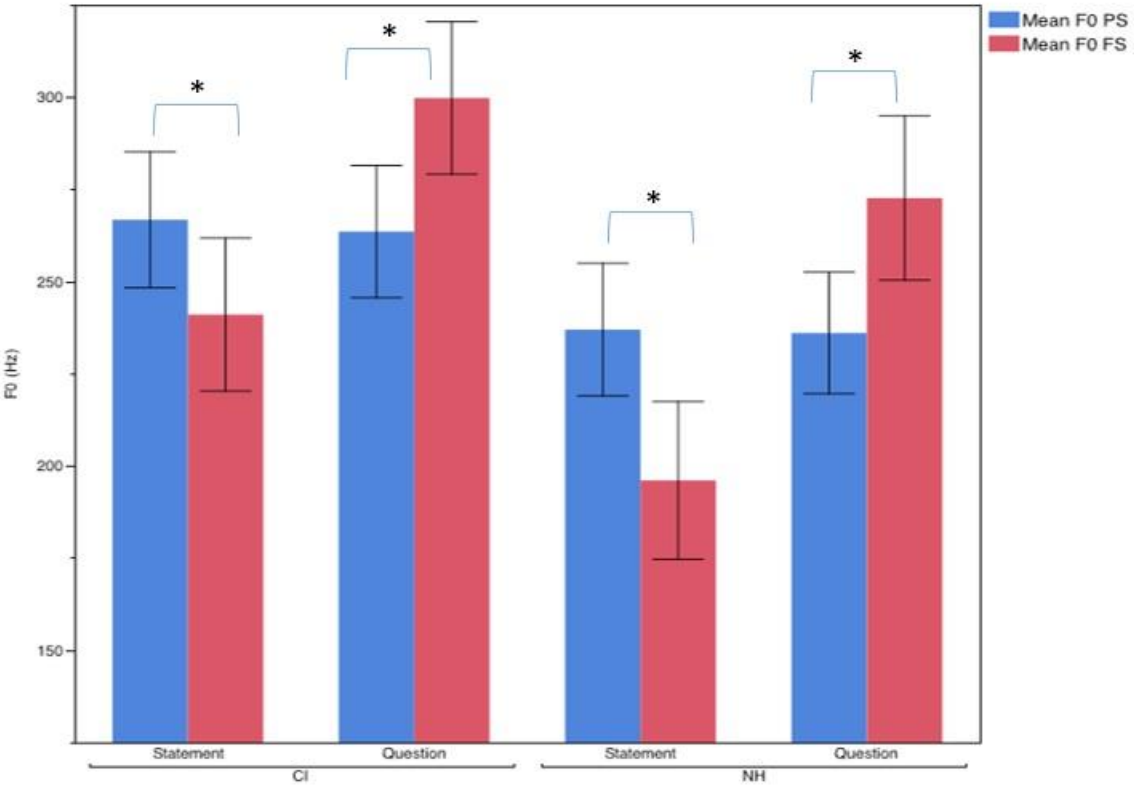

Figure 4. Mean F0 (Hz) measures in PS and FS for elicited questions and statements as produced by CI children and NH children.
Note. Error bars represent 95% confidence intervals. * p < 0.001

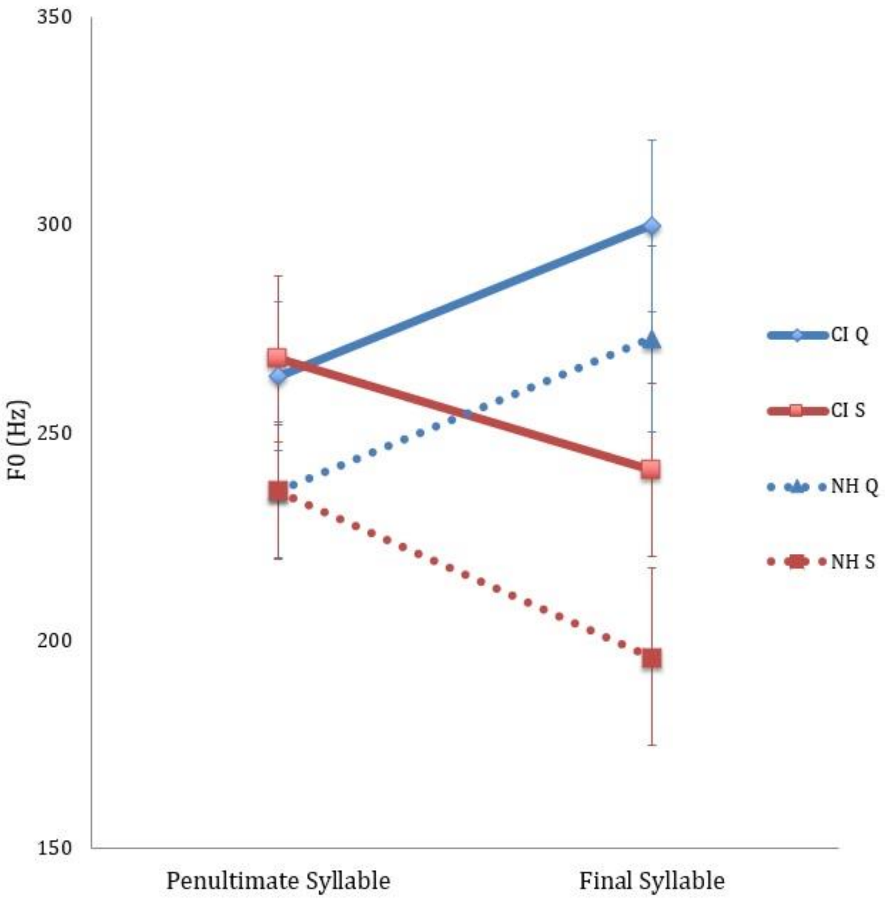

Figure 5. Mean F0 measures in CI and NH participants in the PS and FS. F0 measures are shown to be rising from the PS to the FS in questions, and falling for the same syllables in statements.

Intensity. Shown in Figure 7 are the mean Intensity measures in both the penultimate and final syllables for elicited questions and statements as produced by implanted children and hearing children. The mean intensity for statements by the CI children was 73.93 dB (*SD* = 5.98) for PS and 70.97 dB (*SD* = 7.45) for FS. The mean Intensity for statements by hearing children was 72.93 dB (*SD* = 6.2) for PS and 69.11 dB (*SD* = 6.1) for FS. For questions, the mean Intensity was 73.25 dB (*SD* = 6.22) for PS, and 73.62 dB (*SD* = 5.71) for FS, in implanted children. In hearing children, these measurements were 72.45 dB (*SD* = 5.08) for PS, and 72.77 dB (*SD* = 4.62) for FS.

A three-way mixed ANOVA was conducted to analyze the intensity measures for the two final syllables within the utterances. The between factors were group and elicitation type. The within factor was syllable position.

This ANOVA revealed no main effects for group ($p = 0.121$), elicitation type ($p = 0.074$), or syllable position ($p = 0.08$). However, there was a significant interaction between syllable position and elicitation type ($F_{(1,\ 245)} = 65.432$, $p < 0.0001$). No significant interactions were found between syllable position and group ($p = 0.327$), nor was there a significant three-way interaction among elicitation type, syllable position, and group ($p = 0.379$).

In order to further inspect the interaction between elicitation type and syllable position, a post-hoc analysis using t-tests was conducted, and alpha was adjusted using a Bonferroni correction ($p = 0.0125$). The paired t-tests revealed that for statements, mean intensity for the penultimate syllable was higher than mean intensity for the final syllable both in CWCI ($t_{(1,\ 59)} = 5.897$, $p < 0.0001$) and CWNH ($t_{(1,\ 61)} = 7.882$, $p < 0.0001$). For questions, there was no significant difference in mean intensity between the penultimate and the final syllables for either CWCI ($t_{(1,\ 62)} = -0.897$, $p = 0.373$) or CWNH ($t_{(1,\ 64)} = -0.732$, $p = 0.467$). These results are illustrated in Figure 6. These findings suggest that both CWCI and CWNH decrease intensity from the PS to the FS when producing statements. When producing a question, both groups maintain their mean intensity from the PS to the FS.

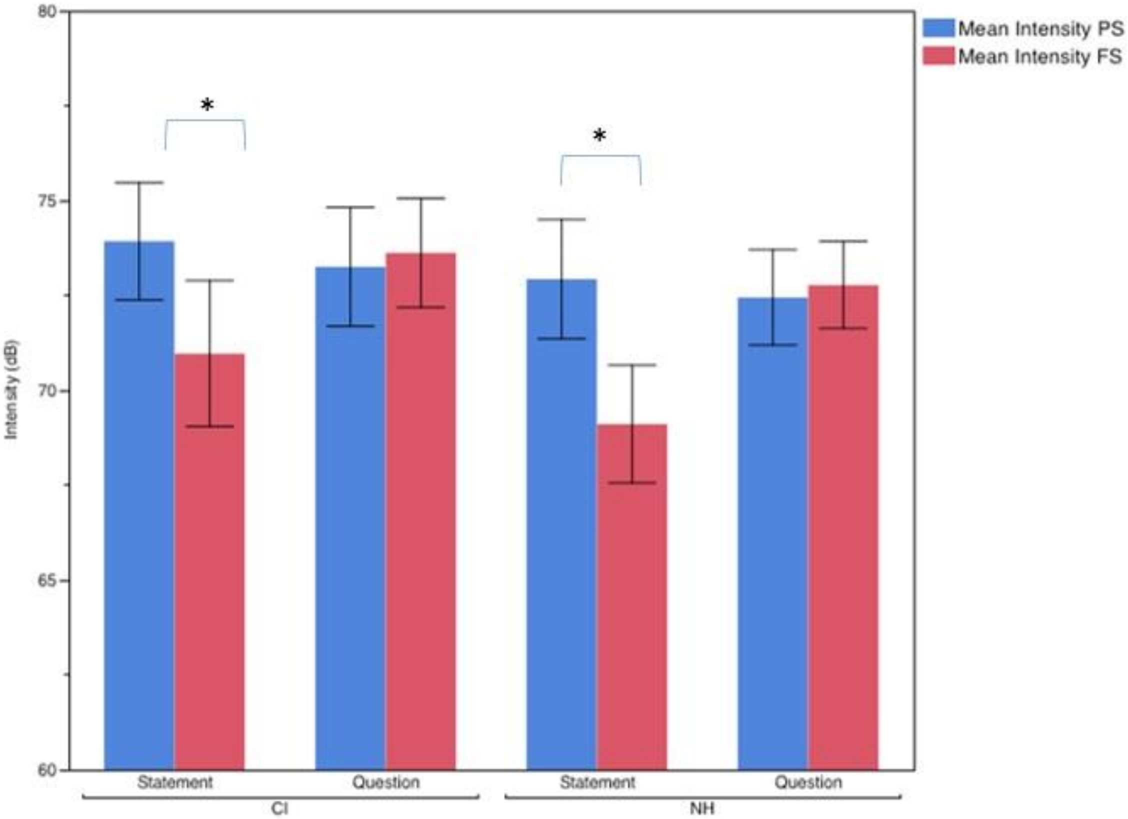

Figure 6. Mean Intensity measures (dB) in PS and FS for elicited questions and statements as produced by CI children and NH children.
Note. Error bars represent 95% confidence intervals. * $p < 0.0001$

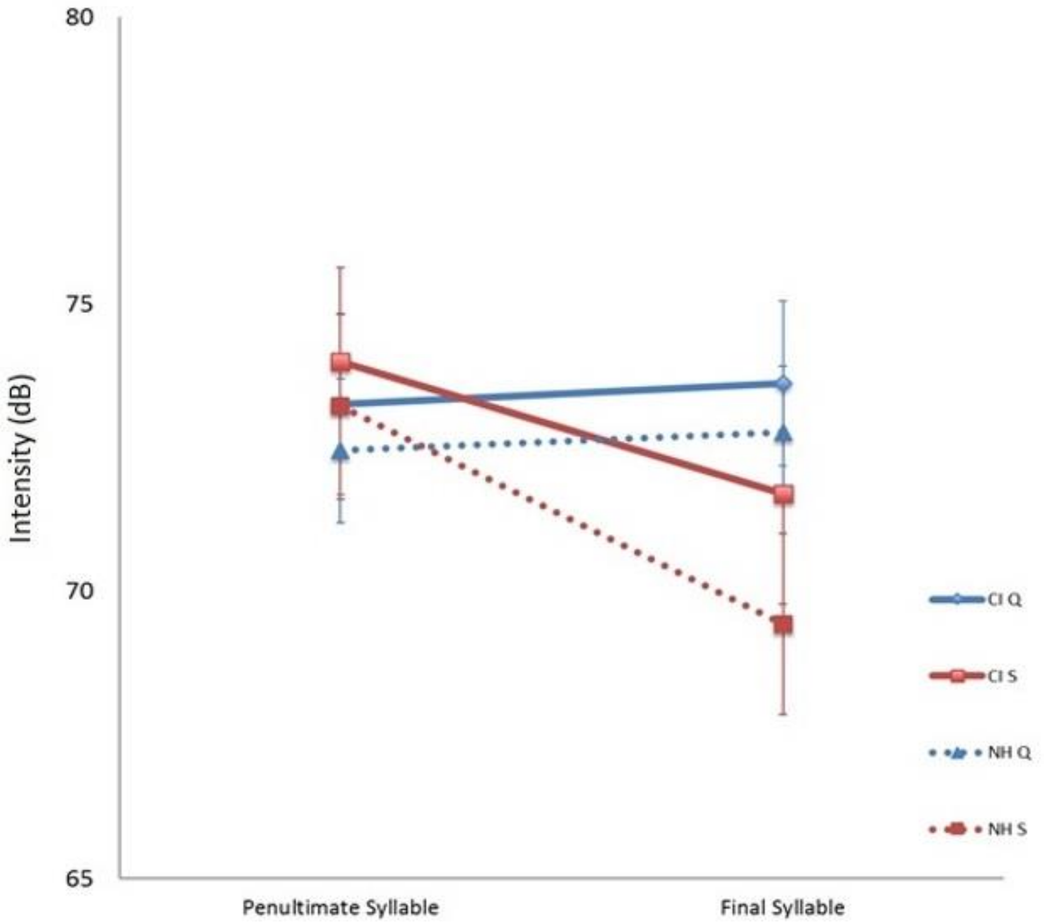

Figure 7. Mean Intensity measures in CI and NH participants in the PS and FS. Mean Intensity measures are shown to be clearly falling for statements from the PS to the FS in both groups. The PS and FS mean Intensity measures for questions appear to remain mostly flat.

Relationship between Listener Judgments and Acoustic Measures

Correlation analyses examined the relation among intonation ratings by listeners and acoustic measures of F0 and intensity. Correlation results are summarized in Table 1. There was a strong[3] positive correlation between listener ratings and mean F0 for the final syllable. There was a moderate positive correlation between listener ratings and mean intensity for the final syllable. There was a strong negative correlation between listener rating and the F0 difference between the penultimate syllable and the final

---

[3] 0.00 – 0.39 = weak correlation
0.40 – 0.59 = moderate correlation
0.60 – 1 = strong correlation
Evans (1996)

syllable. There was a strong negative correlation between listener rating and intensity differences between the penultimate syllable and the final syllable. A moderate positive correlation was noted between mean intensity in the final syllable and mean F0 measures in the final syllable.

Table 1

Correlations between intonation ratings by listeners and acoustic measures

	Mean F0 Penultimate	Mean F0 Final	Mean Intensity Penultimate	Mean Intensity Final	F0 Difference bw Penultimate and Final Syllables	Intensity Difference bw Penultimate and Final Syllables
Mean Intensity Final	.408**	.466*	.809**			
F0 Difference bw Penultimate and Final syllables	.076	-.781*	.149	-.242		
Intensity Difference bw Penultimate and Final syllables	-.003	-.536*	.239	-.378*	.633*	
Listener Rating	.131	.686*	-.023	.385*	-.735*	-.672*

Note. ** $p < 0.0001$; * $p < 0.01$

The relationships between listener judgments and the acoustic measurements (mean F0 & Intensity in the FS) revealed that adult listener's ratings tended to vary with certain acoustic characteristics of the children's productions. For instance, a child rated by a listener as producing a "falling" intonation pattern in an utterance would likely have had acoustic measures that indicated a lower F0 (Hz) and a somewhat low intensity (dB) in the FS. Listener ratings showed a stronger correlation with mean F0 measures for the FS, than with mean intensity measures for the FS. However, the intensity and F0 difference measures also showed a strong correlation with listener judgments. It should be noted that the results from listener ratings and F0 and intensity differences between the penultimate syllables and the final syllables were negative based on the direction of the subtraction (e.g. PS values minus FS values vs. FS values minus PS values). Nevertheless, the strong correlation was indicative of the fact that without a consideration of elicitation type, the change over time from the penultimate syllable to the final syllable in the mean F0 and mean intensity measures were likely used as acoustic cues by the listener judges. Finally, the relationship between the mean Intensity values in the FS and the mean F0 values in the FS revealed that the values generally followed a similar pattern (e.g. a low F0 value in the FS would correspond with a low intensity value in the FS).

No significant correlations were found between the listener ratings and acoustic measurements for F0 and intensity in the PS (refer to Table 1). Moderate positive correlations were noted between the mean F0 measures in the PS and the FS, and mean intensity in the FS. These results indicated that as mean F0 values changed within the penultimate syllable and the final syllable, so did intensity.

Chapter 4: Discussion

The aim of the present study was to characterize the intonation patterns produced by children with cochlear implants and children with normal hearing development, using listener ratings and acoustic measures. Elicited questions and statements allowed for these analyses as well as a closer examination of the relationship between listener ratings and acoustic measures.

The initial research question addressed by this thesis was whether or not adult hearing listeners could distinguish productions of elicited questions and statements in implanted children and normal hearing children. Our results indicated that CWCI and CWNH could produce an intonation distinction as perceived by listeners in elicited questions and statements. When we compared results of the listener ratings of CWCI with results of the listener ratings of CWNH, we found that overall, ratings followed a similar pattern for both groups with greater values for questions, and lower values for statements on the visual analog scale.

Our second question addressed acoustic measurements, asking how CWCI and CWNH controlled F0 and intensity in order to produce distinctions between questions and statements. Mean F0 findings indicated that for both groups, the measures in their final syllables differed, with higher mean F0 in questions than statements. Interestingly, CWCI produced overall higher mean F0 for questions and statements in both the penultimate and final syllables, as compared to the CWNH. These results may be indicative of the effect of the shifted frequency mapping present in CIs. Within the cochlea, a CI typically only stimulates the higher frequencies as the electrodes do not reach the apex, where lower frequencies are stimulable (Nittrouer, 2010).

Mean intensity findings indicated that there were no differences between acoustic measures in the penultimate and the final syllables in questions. This indicated the children maintained the loudness in the final two syllables for elicited questions. For statements, both groups decreased loudness from the penultimate syllable to the final syllable. The intensity measures for statements appear to follow pitch measures, but for questions it does not. And although not significant, overall intensity measures were generally higher in implanted children than in normal hearing children, which are in line with mean F0 measures. These findings are in line with the physiology of falling and rising pitch production. Generally, research has indicated that greater motoric demands are imposed on rising pitch contours as opposed to falling pitch contours (Xu & Sun, 2002). Moreover, research has shown that declination for statements is achieved through reduced F0 and intensity that are the result of falling subglottal pressure at the end of a sentence that affects both F0 and intensity (Lieberman, 1967). According to Lieberman (1967), production of a terminal rising pitch, however, is the result of laryngeal adjustments that are not directly related to subglottal pressure changes that would affect intensity.

In examining the relationship between acoustic cues and listener judgments of questions and statements, the results from the correlation analysis showed that the mean F0 values for the final syllables were strongly related to the listener judgments, more so than the mean intensity measure for the final syllables. Also when removing elicitation type as a factor, both the F0 and Intensity differences between the penultimate syllables and the final syllables were strongly related to listener ratings. This indicated that perhaps the intensity and F0 change over time between the two syllables were used as cues by the adult listeners. This further suggests that listeners focus in on the very last portion or syllable of an utterance when identifying yes/no questions and statements, paying attention to both F0 and intensity when these cues are both present in a speech sample.

Taken together, these findings suggest that both children with early implantation cochlear implants and children with normal hearing are able to produce a distinction between two very different pitch patterns. Although CWCI have limited input from their devices for F0 and to a lesser extent, intensity, they are capable of producing appropriate intonation patterns for questions and statements. Furthermore, normal hearing adult listeners can differentiate between these productions, by rating elicited questions with a higher score than statements on a visual analog scale. These findings are somewhat in line with our outcome predictions. As we anticipated, both groups produced discernible questions and statements; however, CWCI's productions had overall higher pitch and intensity. Contrary to expectation, CWCI appeared similar to hearing children in their use of F0 and intensity to signal questions and statements. Lastly, age related statistical analyses revealed that there were no relationships between age and any F0 or intensity derived measures. While variations within the individual measures of F0 and intensity were present, age did not appear to be a factor.

The results reported here are in partial agreement with previous research. For instance, the study on children with hearing aids by Allen & Arndorfer (2000), reported similar findings that indicated aided children produced a distinction between questions and statements, and their productions were also judged as different by the normal hearing listeners. However, Allen and Arndorfer reported that for the production task, the distinctions by the hearing impaired children were not as clear as the distinctions between questions and statements produced by normal hearing children. Unlike our current study, the participants in the production and perception tasks included older children (7;9 to 14;7), with only 6 hearing impaired children and 6 normal hearing children.

In another study analyzing production and perception of speech intonation that used utterances elicited from pictures, Peng et al. (2008) found that CI children did not demonstrate mastery of speech intonation in both production and perception. Unlike these findings, our study demonstrated that productions of questions and statements by children with cochlear implants were similar to those of normal hearing children and demonstrated; that is, both groups of children produced appropriate rising and falling intonation to mark questions and statements. Differences between Peng et al.'s findings and ours may have been related to sample characteristics (e.g. Peng et al. recorded older children many of whom were implanted at a later age), and elicitation methods (e.g. picture based elicitations vs. role play based elicitations). Age of implantation has been shown to have a significant impact on spoken language outcomes (Connor, 2006; Geers, 2002; Miyamoto, 1994). Because our elicitation technique was designed to examine younger children for whom reading was not an option, elicited conversational utterances were employed. This elicitation may better represent the production capability of the children than a reading task.

Other possible explanations for our findings might lie in areas that were not examined in this thesis. Prosodic cues not addressed in the present study include duration, which has been considered an important factor in marking pitch tracings for questions and statements (Allen & Arndorfer, 2000; Snow, 1994). Lexical or syntactic elements of the utterances could have also played an important factor in listener ratings. In the present study, listeners were instructed to not pay attention to lexical or syntactic information in making their judgments. A recent study comparing intonation judgments of low pass filtered utterances (in which lexical and syntactic information was removed) and intact versions of the same utterances revealed no significant difference in the ratings assigned (Dewey, et al., 2014). In other words, it appears that the presence or absence of lexical and syntactic information was not a factor in the judgments made by adult listeners.

Finally, the results may have been affected by an unexpected aspect of children's productions of connected speech – creaky voice. Creaky voice or glottal fry is phonation type in which the vocal folds are adducted but slack resulting in irregular pulses. Creaky voice usually occurs with lower pitch, but not always. Low frequency productions in questions and statements were often present within the data set (e.g. syllables with frequency < 75 Hz). These low productions in the FS may have influenced F0 measures (Keating, Garellek, & Kreiman, 2015) and resulted in elimination of a small number of productions.

Our findings have helped advance knowledge about productions of elicited questions and statements of implanted children. These findings indicated that CWCI specifically, can accurately produce perceptually distinguishable and acoustically distinguish questions and statements with different pitch patterns. Clinical implications for these findings indicate that typical outcomes for productions of rising and falling pitch patterns (as noted in yes/no questions and statements) can be expected from a young age (3;7) in children with cochlear implants that have had early implantation and have no other disabilities. Other potential clinical implications address the variability in conclusions drawn about CWCI's productions of questions and statements. Our findings suggest that the two groups can accurately produce distinguishable questions and statements in conversation from an age as young as 3;7 years.

Unlike previous studies that tested fewer participants (less than 14 participants in each of the following studies: Lenden & Flipsen, 2006; Allen & Arndorfer, 2000; Patel & Grigos, 2006, Peng, Chatterjee, & Lu, 2014), this study included 14 CI children and 14 NH children. Although the present study examined a greater number of participants than other studies, a total of 28 total participants may still be considered a limitation as it may not include a representative sample of the target population. Another limitation is the small number of utterances per child (4-5 questions, 4-5 statements), for a total of 248 utterances. Although this number is generally higher than has been reported by other studies (e.g. Peng et al., 2007), it may still affect possible generalizations. Although we controlled for extreme noise within our testing rooms, we did not use a sound booth for the recordings. On several occasions, the researchers were required to record in the participants' homes. These unavoidable lack of control added complications to the analysis or in some cases resulted in an inability to analyze a sample altogether. As a result, the methodology for gathering recordings may have limited statistical power, and the generalizations than can be drawn from these results.

Previous research indicated that three parameters (pitch, duration, and intensity) are traditionally responsible for marking yes-no questions and statements (Cruttenden, 1981). In the present study we were only able to look at two – pitch and intensity, marking this as a limitation. Given the nature of the elicitation task, vocabulary was challenging to control, limiting the ability to compare variants of the same words. However, the elicited productions were more in line with more natural productions of children outside the testing room.

A next step might be to design a study that may optimize retrieval of stimuli and the subsequent acoustic analysis of all three acoustic cues – pitch, duration, and intensity. Traditionally, duration has been an important cue for a rising or falling intonation, as mentioned in many studies (Allen & Arndorfer, 2000; Patel & Grigos, 2006; Peng et al., 2007, 2008, 2014), and would therefore be an important addition to production research on children with implants.

Further investigation of cue weighing in adult listeners of rising and falling pitch pattern ratings is also needed. Our results demonstrated that listeners used acoustic cues such as F0 and intensity in order to make distinctions between productions. Nevertheless, it was unclear whether F0 and intensity measures were the only cues used. A study investigating whether there are differences for listeners in the relative weighing of the cues in perceiving distinctions between yes/no questions and statements for CI and hearing children is warranted.

Obtaining accurate F0 measures from utterances was difficult and marked a limitation with current F0 tracking technology in less than optimal room acoustics. Other limitations may have resulted from the methods of the F0 extraction using Praat. However, these measures were verified using a standardized procedure and parameter settings were changed in order to record true F0 measurements when necessary.

It is also noteworthy that while natural speech through the use of elicited questions and statements is a more realistic picture of what a child may produce outside the testing room, it is not guaranteed to result in the desired stimuli. For instance, during our intonation elicitation task, children often produced ambiguous statements. To control for such variability, we developed definitions of ambiguous utterances, or what may be an acceptable utterance and what would entail an unacceptable utterance.

For this study, multiple pitch and intensity measures were obtained (e.g. min F0, max F0, mean F0, slope F0, min Intensity, max Intensity). However, only mean F0 of the PS and the FS (meanF0_PS, meanF0_FS), and mean Intensity of the PS and the FS (meanINT_PS, meanINT_FS) were used for the present analysis, as these measures appeared to be a sufficient F0 acoustic metric for capturing perceptually salient aspects of intonation (Allen & Arndorfer, 2000; Patel & Grigos, 2006; Peng et al., 2007). Nonetheless, future studies might address F0 values such as minimum, maximum and slope, as well as Intensity values such as minimum and maximum. Future studies might include information about CWCI's perception as well as production much like a few previous studies have done (Peng et al., 2008; Allen & Arndorfer, 2000). It would be informative to compare the children's ability to judge similar productions of questions and statements to their own productions, and even to judgments by adult normal hearing listeners.

Chapter 5: Conclusion

This thesis sought to characterize implanted and normal hearing children's productions of elicited questions and statements through listener judgments and through acoustic measurements. The results from the adult listeners' judgments of intonation indicated that both groups of children were judged to have rising intonation patterns for questions, and falling intonation patterns for statements. The acoustic analyses demonstrated that CWCI and CWNH produce distinctively different questions and statements. Further analysis revealed acoustic pitch and intensity measures from the penultimate and final syllables were significantly related to the ratings assigned by listeners. These acoustic measures are in line with previous findings indicating that acoustic cues within the final two syllables of an utterance, and even more so within the final syllable are important for denoting pitch patterns. The findings have implications when aspects of prosody may be considered for speech intervention.

References

Allen, G.D., & Arndorfer, P.M. (2000). Production of sentence-final intonation contours by hearing-impaired children. *Journal of Speech, Language, and Hearing Research*, 43, 441-455.

American Speech-Language-Hearing Association. (2004). *Cochlear implants* [Technical Report]. Available from www.asha.org/policy.

Boersma, P., & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.51, retrieved 21 February 2013 from http://www.praat.org/.

Carrell TC. Level 16 [Computer software] 1998 University of Nebraska; Lincoln, NE: Free download available at http://hush.unl.edu/LabResources.html.

Chatterjee, Y., -R. (1968). Processing fundamental frequency contrasts with cochlear implants: Psychophysics and speech intonation. *Hearing Research*, 235, 145-156.

Connor, C. M., Craig, H. K., Raudenbush, S. W., Heavner, K., & Zwolan, T. A. (2006). The Age at Which Young Deaf Children Receive Cochlear Implants and Their Vocabulary and Speech-Production Growth: Is There an Added Value for Early Implantation? *Ear and Hearing*, 27(6), 628-644. doi:10.1097/01.aud.0000240640.59205.42

Cruttenden, A. (1981). Falls and rises: Meanings and universals. J. Ling. Journal of Linguistics, 17(1), 77. doi:10.1017/s0022226700006782

Crystal, D. (1986). Prosodic development. In P.J. Fletcher and M. Garman (Eds.), *Studies in First Language Development Language Acquisition*, 174-197. doi:10.1017/cbo9780511620683.011

Deroche, M. L., Lu, H., Limb, C. J., Lin, Y., & Chatterjee, M. (2014). Deficits in the pitch sensitivity of cochlear-implanted children speaking English or Mandarin. *Frontiers in Neuroscience*, 8. doi:10.3389/fnins.2014.00282

Fryauf-Bertschy, H., Tyler, R. S., Kelsay, D. M., Gantz, B. J., & Woodworth, G. P. (1997). Implant use by prelingually deafened children: The influence of age at implant and length of device use. *Journal of Speech, Language, and Hearing Research*, 40, 183–199.

Geers, A. E. (2002). Factors affecting the development of speech, language, and literacy in children with early cochlear implantation. *Language, Speech, and Hearing Services in Schools*, 33(3), 172-183.

Geurts, L., & Wouters, J. (2001). Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *Journal of the Acoustical Society of America*, 109, 713-726.

Green, T., Faulkner, A., & Rosen, S. (2004). Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. *The Journal of the Acoustical Society of America* 116(4): 2298-2310.

Hitech (2008) Soundswell. 4.5 ed. Sweden.

Kawahara, S. (2010, 2) *Intensity Script.* Retrieved from http://user.keio.ac.jp/~kawahara/scripts/get_intensity_minmax.praat

Keating, P., Garellek, M., & Kreiman, J. Acoustic properties of different kinds of creaky voice.

Lehiste, I. (1976). Suprasegmental Features of Speech. *Contemporary Issues in Experimental Phonetics*, 225-239. doi:10.1016/b978-0-12-437150-7.50013-0

Lenden, J. M., & Flipsen, P. Jr. (2007). Prosody and voice characteristics of children with cochlear implants. *Journal of Communication Disorders* 40(1): 66-81.

Lennes, M. (2003, 4). *Pitch Script*. Retrieved from http://www.helsinki.fi/~lennes/praatscripts/public/collect_pitch_data_from_files.praat

Lieberman, P. (1967). *Intonation, perception, and language.* Cambridge: The MIT Press.

Loeb, D. F., Allen, G. D. (1993). Preschoolers' imitation of intonation contours. *J. Speech Hearing Res.* 36, 4-13.

Mahshie, J., Core, C, Choi, S., & Dewey, R. (2014). Development of intonation production for yes/no questions and statements by English speaking children with cochlear implants. *Tone and Intonation in Europe Conference,* September 9-12, Utrecht, NL.

Mahshie, J., Core, C., & Larsen, M. (2015). Auditory Perception and Production of Speech Feature Contrasts by Pediatric Implant Users. *Ear & Hearing*, 36, 653-663.

Miyamoto, R. T., Osberger, M. J., Todd, S. L., Robbins, A. M., Stroer, B. S., Zimmerman-Phillips, S., & Carney, A.E. (1994). Variables affecting implant performance in children. *Laryngoscope*, 104, 1120-1124.

Nittrouer, S. (2010). *Early Development in Children with Hearing Loss*

Patel, R., & Grigos, M.I. (2006). Acoustic characterization of the question-statement contrast in 4, 7 and 11 year-old children. *Speech Communication,* 48, 1308-1318.

Peng, S.C., Tomblin, J.B., Cheung, H., Lin, Y.S., & Wang, L.S. (2004). Perception and production of mandarin tones in prelingually deaf children with cochlear implants. *Ear and Hearing,* 25, 251-264.

Peng, S.C., Tomblin, J.B., Spencer, L.J., & Hurtig, R.R. (2007). Imitative production of rising speech intonation in pediatric cochlear implant recipients. *Journal of Speech, Language, and Hearing Research,* 50, 1210-1227.

Peng, S.C., Tomblin, J.B., & Turner, C.W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear and Hearing*, 29, 336-51.

See, R.L., Driscoll, V.D., Gfeller, K., Kliethermes, S., & Oleson, J. (2013). Speech Intonation and Melodic Contour Recognition in Children with Cochlear Implants and with Normal Hearing. *Otol Neurotol.* 34(3), 490–498. doi:10.1097/MAO.0b013e318287c985

Smith, B. (2013). Segmenting Speech. Retrieved March 18, 2016, from http://www.ling.ohio-state.edu/~bsmith/docs/speech_segmentation_vocabdrill

Shriberg, L.D., & Kent, R. D. (1982). *Clinical Phonetics*. New York: Wiley.

Titze, I. R., & Titze, M. R. (1993). Phonation threshold pressure in a physical model of the vocal folds. *The Journal of the Acoustical Society of America J. Acoust. Soc. Am.*, 94(3), 1762. doi:10.1121/1.408054

Torppa, R., Faulkner, A., Huotilainen, M., Järvikivi, J., Lipsanen, J., Laasonen, M., & Vainio, M. (2014). The perception of prosody and associated auditory cues in early-implanted children: The role of auditory working memory and musical activities. *International Journal of Audiology*, 53(3).

Tserdanelis, G., & Wong, W. Y. P. (2004). Language files. *Materials for an Introduction to Language and Linguistics*, 531.

Tyler, R. S., Teagle, H. F. B., Kelsay, D. M. R., Gantz, B. J., Woodworth, G. G., & Parkinson, A. J. (2000). Speech perception by prelingually deaf children after six years of cochlear implant use: Effects of age at implantation. *Annals of Otology, Rhinology, & Laryngology*, 109, 82–84.

Wright, R., & Nichols, D. (2015, June 18). Measuring Vowel Duration. Retrieved March 19, 2016, from https://zeos.ling.washington.edu/~labwiki/w/index.php?title=Measuring_Vowel_Duration

Xu, Y., Sun, X. (2002). Maximum speed of pitch change and how it may relate to speech. *J. Acoust. Soc. Amer.* 111, 1399-1413.

Yvan, R., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O'Brien, P., & Wareham, T., (2006). Introducing Phon: A Software Solution for the Study of Phonological Acquisition. In David Bamman, Tatiana Magnitskaia & Colleen Zaller (eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. 489-500.

Zhu, M., Fu, Q., Galvin, J. J., Jiang, Y., Xu, J., Xu, C., & Chen, B. (2011). Mandarin Chinese speech recognition by pediatric cochlear implant users. *International Journal of Pediatric Otorhinolaryngology,* 75(6), 793-800. doi:10.1016/j.ijporl.2011.03.009

Appendix A

Table A1.

Age of Participants

CWCI			*CWNH*	
ID	*Age at Testing (months)*	*Age at Implantation (months)*	*ID*	*Age at Testing (months)*
Ci007	43	14	Ci525	40
Ci028	44	15	Ci517	41
Ci029	44	26	Ci530	46
Ci013	47	12	Ci510	54
Ci005	49	35	Ci511	54
Ci014	58	10	Ci518	55
Ci022	58	11	Ci522	59
Ci023	61	29	Pv502	61
Ci003	64	18	Pv501	62
Ci006	64	18	Pv504	68
Ci004	70	19	Ci503	76
Pv009	75	15	Ci527	76
Ci010	83	16	Ci526	77
Ci015	89	13	Ci521	88
Mean	**60.64**	**17.93**	**Mean**	**61.21**
St.Dev.	**14.20**	**7.00**	**St.Dev.**	**13.86**

Note. A description of the ages of each matched pair of participants, as well as the individual age at implantation is provided. The mean and standard deviation for each group of children are also shown.

Table A2.

Description of CWCI

ID	Age of Activation R/L	CI Type	Duration of CI Use at Testing R/L	Use of HA Before CI
Ci007	1;4/HA	Cochlear Internal Freed./External 5	2;4/HA	Yes
Ci028	1;4/1;10	Cochlear Nucleus 5	2;3/1;6	Yes
Ci029	Unknown/2;4	Cochlear External N5	Unknown/1;5	Yes
Ci013	1;1/1;7	Cochlear Freedom	2;11/2;5	Yes
Ci005	1;1/3;1	Cochlear Freedom BTE Mini	3;0/1;0	Yes
Ci014	0;10/3;9	Advanced Bionics Harmony	3;11/1;1	Yes
Ci022	0;11/0;11	Cochlear	5;0/5;0	Yes
Ci023	HA/unknown	Cochlear Nucleus Freedom	HA/unknown	Yes
Ci003	Unknown/1;7	Cochlear Freedom BTE	Unknown/3;9	Yes
Ci006	Unknown/1;7	Cochlear Freedom	Unknown/3;9	Yes
Ci004	5;2/1;9	Advanced Bionics Harmony	0;9/4;6	Yes
Pv009	1;1/1;6	AB 90k Body Worn	2;6/2;3	Yes
Ci010	1;5/6;3	CP 810 Nucleus 5 (R) Cochlear Freedom (L)	5;6/0;8	Unknown
Ci015	7;1/1;4	Nucleus 5	0;4/6;1	Yes

Note. Information regarding the age of activation for the right and left ears for the CIs, the type of CI device, the experience with the CI device by the date of testing, and the use of hearing aids before CI surgery is provided.

Appendix B: List of stimuli, including transcriptions, elicitation type

Group	Age	Subject ID	Record Number	Elicitation	Orthography
CI	64	Ci003	1	S	Hi buddy.
CI	64	Ci003	12	S	I like carrots.
CI	64	Ci003	8	S	And he ate an apple.
CI	64	Ci003	14	Q	Do you want an apple?
CI	64	Ci003	15	Q	Do you want anything else?
CI	64	Ci003	5	Q	You want some cake?
CI	64	Ci003	6	Q	Do you want a cracker?
CI	70	Ci004	13	S	You want milk.
CI	70	Ci004	14	S	Don’t know.
CI	70	Ci004	6	S	Strawberry.
CI	70	Ci004	7	S	Carrot.
CI	70	Ci004	11	Q	You want a cracker?
CI	70	Ci004	2	Q	Are you hungry?
CI	70	Ci004	3	Q	Get the knife ?
CI	70	Ci004	5	Q	You want to keep?
CI	70	Ci004	8	Q	You want a carrot?
CI	49	Ci005	11	S	A bed.
CI	49	Ci005	12	S	Blanket.
CI	49	Ci005	6	S	A carrot.
CI	49	Ci005	1	Q	Are you hungry?
CI	49	Ci005	10	Q	Do you want a blanket?
CI	49	Ci005	2	Q	Do you want a strawberry?
CI	49	Ci005	3	Q	Do you want another strawberry?
CI	49	Ci005	9	Q	Are you sleepy?
CI	64	Ci006	12	S	Here's one.
CI	64	Ci006	4	S	Thank you.
CI	64	Ci006	5	S	I have pretzels.
CI	64	Ci006	7	S	I can give you one.
CI	64	Ci006	10	Q	Do you want another strawberry?
CI	64	Ci006	3	Q	Would you like something to eat?
CI	64	Ci006	6	Q	Do you like pretzels?
CI	64	Ci006	8	Q	Do you like strawberries?
CI	43	Ci007	1	S	Let me try.
CI	43	Ci007	2	S	Piece of cake right on here.
CI	43	Ci007	8	S	He can have some pretzel.
CI	43	Ci007	9	S	Another cake.
CI	43	Ci007	10	Q	You want carrot?

CI	43	Ci007	3	Q	You want a piece of cake?
CI	43	Ci007	5	Q	You want milk?
CI	43	Ci007	6	Q	You want pretzel?
CI	74	Pv009	19	S	I like him
CI	74	Pv009	22	S	I like cheese.
CI	74	Pv009	5	S	I like jam
CI	74	Pv009	6	S	I like carrots
CI	74	Pv009	1	Q	Do you want to still eat a carrot?
CI	74	Pv009	12	Q	Do you like cupcake?
CI	74	Pv009	18	Q	Do you like apples?
CI	74	Pv009	21	Q	Want some cheese?
CI	74	Pv009	3	Q	Do you want a carrot?
CI	74	Pv009	7	Q	Do you want some lettuce?
CI	83	Ci010	11	S	Yogurt.
CI	83	Ci010	12	S	Okay.
CI	83	Ci010	13	S	Only these two.
CI	83	Ci010	3	S	Hi Bo.
CI	83	Ci010	6	S	Sorry.
CI	83	Ci010	10	Q	Can I have a cake?
CI	83	Ci010	14	Q	Do you like pretzels?
CI	83	Ci010	7	Q	Do you wanna have a strawberry?
CI	83	Ci010	8	Q	Do you wanna have this?
CI	47	Ci013	1	S	Don't know.
CI	47	Ci013	10	S	Four candles for his birthday to blow them out.
CI	47	Ci013	11	S	And the birthday cake.
CI	47	Ci013	12	S	Some more cake.
CI	47	Ci013	4	S	That sounds fun.
CI	47	Ci013	3	Q	Do you want some cake?
CI	47	Ci013	5	Q	Do you want a birthday candle?
CI	47	Ci013	7	Q	Would you like some candles for your birthday?
CI	47	Ci013	8	Q	Do you want some more candles?
CI	47	Ci013	9	Q	Would you like some pretzel?
CI	58	Ci014	10	S	Do you want a knife?
CI	58	Ci014	11	S	Put it in the oven
CI	58	Ci014	4	S	I don’t know.
CI	58	Ci014	8	S	Peanut butter.
CI	58	Ci014	1	Q	You want apple?
CI	58	Ci014	2	Q	You like carrots?
CI	58	Ci014	3	Q	Crackers?
CI	58	Ci014	5	Q	You want it?
CI	58	Ci014	6	Q	Want cupcake?

CI	89	Ci015	10	S	This one knife.
CI	89	Ci015	6	S	Strawberry.
CI	89	Ci015	7	S	A strawberry cake.
CI	89	Ci015	9	S	Celery.
CI	89	Ci015	2	Q	Do you wanna eat crackers?
CI	89	Ci015	4	Q	Do you want any something else?
CI	89	Ci015	5	Q	Do you want some pretzels Bo?
CI	58	Ci022	3	S	You want some red XX[4]?
CI	58	Ci022	4	S	Chocolate.
CI	58	Ci022	5	S	I like chocolate cake.
CI	58	Ci022	7	S	Yogurt.
CI	58	Ci022	8	S	Strawberry.
CI	58	Ci022	13	Q	Bo do you want milk?
CI	58	Ci022	16	Q	Bo do you want some bread?
CI	58	Ci022	17	Q	Did you want a sandwich?
CI	58	Ci022	19	Q	Do you want some grapes?
CI	58	Ci022	6	Q	Bo do you want some cake?
CI	61	Ci023	1	S	You want to eat strawberry.
CI	61	Ci023	11	S	Strawberry cake.
CI	61	Ci023	3	S	Carrot.
CI	61	Ci023	7	S	This one.
CI	61	Ci023	8	S	Strawberry.
CI	61	Ci023	10	Q	You want sandwich?
CI	61	Ci023	2	Q	You want strawberry?
CI	61	Ci023	4	Q	You want carrots?
CI	61	Ci023	5	Q	You want cracker?
CI	61	Ci023	6	Q	You want crackers?
CI	61	Ci023	9	Q	You want pretzel, doggie?
CI	44	Ci028	1	S	Um, strawberry (strawberr).
CI	44	Ci028	10	S	I wanna eat all of it.
CI	44	Ci028	14	S	Strawberry (strawberr).
CI	44	Ci028	18	S	Apple.
CI	44	Ci028	21	S	It's a carrot.
CI	44	Ci028	13	Q	Do you want some?
CI	44	Ci028	15	Q	You want apple?
CI	44	Ci028	17	Q	You want apple?
CI	44	Ci028	3	Q	You want some?
CI	44	Ci028	5	Q	You want some?
CI	44	Ci029	10	S	Here's a carrot.

---

[4] XX = unintelligible

CI	44	Ci029	12	S	Do you have strawberry?
CI	44	Ci029	14	S	This is the apple.
CI	44	Ci029	2	S	Do you want some milk?
CI	44	Ci029	11	Q	Want carrots?
CI	44	Ci029	5	Q	Want milk?
CI	44	Ci029	7	Q	Unint - 3 syllables.
NH	62	Pv501	10	S	Apple.
NH	62	Pv501	3	S	Yogurt.
NH	62	Pv501	4	S	Strawberry yogurt.
NH	62	Pv501	14	Q	You want strawberry cake?
NH	62	Pv501	20	Q	You want some nice strawberries?
NH	62	Pv501	5	Q	You want yogurt?
NH	61	Pv502	14	S	String cheese.
NH	61	Pv502	15	S	A sandwich (samwich).
NH	61	Pv502	17	S	I bought the carrot on the celery.
NH	61	Pv502	19	S	Put peanut butter on it and raisin.
NH	61	Pv502	6	S	Strawberry yogurt.
NH	61	Pv502	10	Q	Do you want an apple?
NH	61	Pv502	16	Q	Did you want a sandwich?
NH	61	Pv502	20	Q	Do you want a carrot on the celery?
NH	61	Pv502	4	Q	Do you want a strawberry?
NH	61	Pv502	8	Q	Do you want yogurt?
NH	76	Ci503	12	S	Bo, do you want some celery?
NH	76	Ci503	3	S	Do you want the yogurt?
NH	76	Ci503	9	S	Chocolate.
NH	76	Ci503	5	Q	You want some milk?
NH	76	Ci503	6	Q	Do you want some strawberry?
NH	76	Ci503	7	Q	Do you want the carrot?
NH	77	Ci509	1	S	Strawberry.
NH	77	Ci509	23	S	Here's some apple.
NH	77	Ci509	24	S	Maybe cracker and some of this.
NH	77	Ci509	28	S	And that's it.
NH	77	Ci509	10	Q	Would you like some cake?
NH	77	Ci509	2	Q	Do you want some strawberries?
NH	77	Ci509	4	Q	Do you want grapes?
NH	77	Ci509	6	Q	Do you want some bread?
NH	77	Ci509	9	Q	Do you want apples?
NH	54	Ci510	1	S	Hi Bo.
NH	54	Ci510	5	S	Inside this is the carrot.
NH	54	Ci510	7	S	Blueberry.
NH	54	Ci510	14	Q	Do you want celery?

NH	54	Ci510	16	Q	Do you want a carrot?
NH	54	Ci510	18	Q	Do you want a strawberry?
NH	54	Ci510	22	Q	You want a apple?
NH	54	Ci510	3	Q	He have milk?
NH	54	Ci511	1	S	Cookie.
NH	54	Ci511	3	S	A carrot.
NH	54	Ci511	5	S	Strawberry.
NH	54	Ci511	7	S	Yogurt.
NH	54	Ci511	8	S	Strawberry yogurt.
NH	54	Ci511	15	Q	Do you want a carrot?
NH	54	Ci511	17	Q	Do you want a pretzel?
NH	54	Ci511	2	Q	Do you want a cookie?
NH	54	Ci511	6	Q	Do you want strawberry cake?
NH	54	Ci511	9	Q	Do you want strawberry yogurt?
NH	41	Ci517	1	S	Tomato.
NH	41	Ci517	14	S	Bow tie noodle.
NH	41	Ci517	5	S	A sandwich.
NH	41	Ci517	9	S	Pineapple.
NH	41	Ci517	15	Q	Bo would you like to um bow tie noodles?
NH	41	Ci517	20	Q	Bo would you like some ham?
NH	41	Ci517	21	Q	Bo would you like some ham?
NH	41	Ci517	3	Q	Bo would you like tomato?
NH	55	Ci518	19	S	Sandwich.
NH	55	Ci518	25	S	Cupcake.
NH	55	Ci518	4	S	Sausage.
NH	55	Ci518	6	S	I don’t know.
NH	55	Ci518	9	S	Carrot.
NH	55	Ci518	10	Q	Bo do you want carrots?
NH	55	Ci518	12	Q	Bo do you like cheese?
NH	55	Ci518	17	Q	Bo, would you like a sandwich?
NH	55	Ci518	21	Q	Bo, do you like spinach?
NH	55	Ci518	8	Q	Bo would you like pasta?
NH	68	Pv520	18	S	I would like a tomato.
NH	68	Pv520	23	S	I would like a pineapple.
NH	68	Pv520	2	S	I would like a cupcake.
NH	68	Pv520	5	S	I don't like the eggplant.
NH	68	Pv520	7	S	I would like a chocolate cupcake.
NH	68	Pv520	13	Q	Would you like this fish?
NH	68	Pv520	22	Q	Would you like some corn?
NH	68	Pv520	24	Q	Would you like some pasta?
NH	68	Pv520	32	Q	Would you like an apple?

NH	68	Pv520	8	Q	Do you want the lettuce?
NH	88	Ci521	1	Q	Hello.
NH	88	Ci521	18	S	Mushroom.
NH	88	Ci521	23	S	Green pepper.
NH	88	Ci521	14	Q	The cheese.
NH	88	Ci521	31	Q	No icing cake.
NH	88	Ci521	12	S	Do you want a pineapple?
NH	88	Ci521	15	S	Do you want some cheese?
NH	88	Ci521	3	S	Do you want some bread?
NH	88	Ci521	10	Q	Do you want a cupcake?
NH	88	Ci521	7	Q	Do you want a plum?
NH	59	Ci522	10	S	Orange.
NH	59	Ci522	11	S	Tomato.
NH	59	Ci522	13	S	Lettuce.
NH	59	Ci522	17	S	A flower cake.
NH	59	Ci522	6	S	Carrot.
NH	59	Ci522	12	Q	Do you want tomato?
NH	59	Ci522	14	Q	Do you want lettuce?
NH	59	Ci522	5	Q	Do you want this?
NH	59	Ci522	7	Q	Do you want this?
NH	59	Ci522	8	Q	Do you want carrots?
NH	40	Ci525	10	S	Cupcake.
NH	40	Ci525	13	S	Strawberry.
NH	40	Ci525	16	S	I'm gonna give it to him.
NH	40	Ci525	2	S	Pizza.
NH	40	Ci525	6	S	Apple.
NH	40	Ci525	14	Q	Do you want cupcake?
NH	40	Ci525	23	Q	You don't want it?
NH	40	Ci525	3	Q	Do you want it?
NH	40	Ci525	5	Q	You want pineapple?
NH	76	Ci527	10	S	Chocolate.
NH	76	Ci527	18	S	Oh, lettuce.
NH	76	Ci527	5	S	Hot dog.
NH	76	Ci527	9	S	Slice the cake in half.
NH	76	Ci527	13	Q	Do you want the whole thing?
NH	76	Ci527	15	Q	You want carrots?
NH	76	Ci527	20	Q	Does he gotta eat all of them?
NH	76	Ci527	3	Q	You want a mushroom?
NH	76	Ci527	4	Q	Do you want a hot dog?
NH	46	Ci530	17	S	I don’t like pineapple.
NH	46	Ci530	22	S	I like noodles.

NH	46	Ci530	2	S	He wants to eat a watermelon.
NH	46	Ci530	38	S	I like noodles.
NH	46	Ci530	9	S	Cause I like them at a party.
NH	46	Ci530	11	Q	Do you want some cake?
NH	46	Ci530	28	Q	You want another hotdog?
NH	46	Ci530	30	Q	Do you want some fish?
NH	46	Ci530	5	Q	Do you want a carrot?
NH	46	Ci530	7	Q	Do you want a cupcake?

Appendix C: Modified Pitch Script

```
form Analyze pitch maxima from labeled segments in files
	comment Directory of sound files
	text sound_directory C:\Users\X\
	sentence Sound_file_extension .wav
	comment Directory of TextGrid files
	text textGrid_directory C:\Users\X\
	sentence TextGrid_file_extension .TextGrid
	comment Full path of the resulting text file:
	text resultfile C:\Users\X\Pitch Results\Pitch Results.txt
	text resultfile2 C:\Users\X\Pitch Results\Pitch Results2.txt
	text resultfile3 C:\Users\X\Pitch Results\Pitch Results3.txt
	comment Which tier do you want to analyze?
	sentence Tier VoicedSegment
	comment Pitch analysis parameters
	positive my_step 0.01
# Time_step = 0.0 => Time_step = 0.75 / Minimum_pitch_ (Pitch floor)

endform

# Here, you make a listing of all the sound files in a directory.
# The example gets file names ending with ".wav" from D:\tmp\

Create Strings as file list... list 'sound_directory$'*'sound_file_extension$'
numberOfFiles = Get number of strings

# Check if the result file exists:
if fileReadable (resultfile2$)
	pause The result file 'resultfile2$' already exists! Do you want to overwrite it?
	filedelete 'resultfile2$'
endif

# Write a row with column titles to the result file:
# (remember to edit this if you add or change the analyses!)

titleline2$ = "filename	segmentlabel	startTime	endTime	PS_f0min	PS_tmin
	PS_f0max	PS_tmax	PS_meanf0	PS_f0range_oct
	PS_f0minmax_timedistance	PS_slope	PS_f0range_sts'newline$'"
fileappend "'resultfile2$'" 'titleline2$'

# Check if the result file exists:
if fileReadable (resultfile3$)
	pause The result file 'resultfile3$' already exists! Do you want to overwrite it?
	filedelete 'resultfile3$'
endif

# Write a row with column titles to the result file:
# (remember to edit this if you add or change the analyses!)

titleline3$ = "filename	segmentlabel	startTime	endTime	FS_f0min	FS_tmin
	FS_f0max	FS_tmax	FS_meanf0	FS_f0range_oct
	FS_f0minmax_timedistance	FS_slope	FS_f0range_sts'newline$'"
fileappend "'resultfile3$'" 'titleline3$'
```

```
# Check if the result file exists:
if fileReadable (resultfile$)
    pause The result file 'resultfile$' already exists! Do you want to overwrite it?
    filedelete 'resultfile$'
endif

# Write a row with column titles to the result file:
# (remember to edit this if you add or change the analyses!)

titleline$ = "filename          segmentlabel          time(s)          pitch(Hz)'newline$'"
fileappend "'resultfile$'" 'titleline$'

# Go through all the sound files, one by one:

for ifile to numberOfFiles
    filename$ = Get string... ifile
    # A sound file is opened from the listing:
    Read from file... 'sound_directory$''filename$'
    # Starting from here, you can add everything that should be
    # repeated for every sound file that was opened:
    soundname$ = selected$ ("Sound", 1)
    To Pitch (ac): 0.01, 75, 15, "no", 0.03, 0.45, 0.01, 0.35, 0.14, 600
    # Open a TextGrid by the same name:
    gridfile$ = "'textGrid_directory$''soundname$''textGrid_file_extension$'"
    if fileReadable (gridfile$)
        Read from file... 'gridfile$'
        # Find the tier number that has the label given in the form:
        call GetTier 'tier$' tier
        numberOfIntervals = Get number of intervals... tier
        # Pass through all intervals in the selected tier:
        for interval to numberOfIntervals
            label$ = Get label of interval... tier interval
            if label$ = "PS"
                start = Get starting point... tier interval
                end = Get end point... tier interval
                select Pitch 'soundname$'

                f0min = Get minimum... start end Hertz Parabolic
                tmin = Get time of minimum... start end Hertz Parabolic
                f0max = Get maximum... start end Hertz Parabolic
                tmax = Get time of maximum... start end Hertz Parabolic
                meanf0 = Get mean... start end Hertz
                f0range_oct = log2(f0max) - log2(f0min)
                f0minmax_timedistance = abs(tmax - tmin)
                slope = f0range_oct / f0minmax_timedistance
                f0range_sts = 12*f0range_oct

                resultline2$ = "'soundname$' 'label$' 'start'    'end'     'f0min'   'tmin'
                        'f0max'   'tmax'    'meanf0'   'f0range_oct'
                    'f0minmax_timedistance'   'slope''f0range_sts'   'newline$'"
                fileappend "'resultfile2$'" 'resultline2$'

                    # get the values at that interval
                    for i to (end-start)/my_step
                        time = start + i * my_step
                        pitch = Get value at time: time, "Hertz", "Linear"
```

```
                    printline 'pitch'
                    # Save result to text file
                    resultline$ = "'soundname$'        'label$'    'time'
                                                  'pitch"newline$'"
                    fileappend "'resultfile$'" 'resultline$'

                endfor
                select TextGrid 'soundname$'
        endif


        if label$ = "FS"
            start = Get starting point... tier interval
            end = Get end point... tier interval
            select Pitch 'soundname$'

            f0min = Get minimum... start end Hertz Parabolic
            tmin = Get time of minimum... start end Hertz Parabolic
            f0max = Get maximum... start end Hertz Parabolic
            tmax = Get time of maximum... start end Hertz Parabolic
            meanf0 = Get mean... start end Hertz
            f0range_oct = log2(f0max) - log2(f0min)
            f0minmax_timedistance = abs(tmax - tmin)
            slope = f0range_oct / f0minmax_timedistance
            f0range_sts = 12*f0range_oct

            resultline3$ = "'soundname$ 'label$' 'start'    'end'      'f0min'   'tmin'
                                'f0max'   'tmax'    'meanf0'   'f0range_oct'
                  'f0minmax_timedistance'   'slope"f0range_sts'   'newline$'"
            fileappend "'resultfile3$'" 'resultline3$'

                # get the values at that interval
                for i to (end-start)/my_step
                    time = start + i * my_step
                    pitch = Get value at time: time, "Hertz", "Linear"
                    printline 'pitch'
                    # Save result to text file
                    resultline$ = "'soundname$'        'label$'    'time'
                                                  'pitch"newline$'"
                    fileappend "'resultfile$'" 'resultline$'

                endfor
                select TextGrid 'soundname$'
        endif

    endfor
    # Remove the TextGrid object from the object list
    select TextGrid 'soundname$'
    Remove
endif
# Remove the temporary objects from the object list
select Sound 'soundname$'
plus Pitch 'soundname$'
Remove
select Strings list
# and go on with the next sound file!
```

```
endfor

Remove

#-------------
# This procedure finds the number of a tier that has a given label.

procedure GetTier name$ variable$
        numberOfTiers = Get number of tiers
        itier = 1
        repeat
                tier$ = Get tier name... itier
                itier = itier + 1
        until tier$ = name$ or itier > numberOfTiers
        if tier$ <> name$
                'variable$' = 0
        else
                'variable$' = itier - 1
        endif

              if 'variable$' = 0
                            exit The tier called 'name$' is missing from the file 'soundname$'!
              endif

endproc
```

Appendix D: Modified Intensity Script

```
# This Praat script will get average intensity, minimal intensity, and maximal intensity (in dB) of all labeled intervals of all (or a specified set of) files in a folder.
# To use, you specifiy a folder with wav.files, and base names if you want to analyze only a subset of files.
# The script assumes that you already have labeled intervals. The textgrid files and sound files should have the same name.

form Get Intensity
    sentence Directory C:/Users/X/
    comment If you want to analyze all the files, leave this blank
    word Base_file_name
    comment The name of result file
    text textfile C:/Users/X/intensity_list.txt
    comment Which tier do you want to analyze?
    sentence Tier VoicedSegment
endform

#Print one set of headers

fileappend "'textfile$'" File name'tab$'Interval name'tab$'Avg Int'tab$'Min Int'tab$'Min Int Time'tab$'Max Int'tab$'Max Int time'tab$'
fileappend "'textfile$'" 'newline$'

#Read all files in a folder

Create Strings as file list... wavlist 'directory$'/'base_file_name$'*.wav
Create Strings as file list... gridlist 'directory$'/'base_file_name$'*.TextGrid
n = Get number of strings


for i to n
clearinfo
#We first extract intensity tiers
    select Strings wavlist
    filename$ = Get string... i
    Read from file... 'directory$'/'filename$'
    soundname$ = selected$ ("Sound")
    To Intensity... 100 0

# We print out the file names
    labelline$ = "'soundname$''tab$'"
    fileappend "'textfile$'" 'labelline$'

# We now read grid files and extract all intervals in them
    select Strings gridlist
    gridname$ = Get string... i
    Read from file... 'directory$'/'gridname$'
    # Find the tier number that has the label given in the form:
    call GetTier 'tier$' tier
    int=Get number of intervals... tier

# We calculate intensity for all labeled intervals
for k from 1 to 'int'
    select TextGrid 'soundname$'
```

```
        label$ = Get label of interval... tier 'k'
        if label$ = "PS"[5]

                # calculates the onset and offset
                onset = Get starting point... tier 'k'
                offset = Get end point... tier 'k'

                #calculates the intensity values
                select Intensity 'soundname$'
                min_int = Get minimum... onset offset Parabolic
                min_time = Get time of minimum... onset offset Parabolic
                max_int = Get maximum... onset offset Parabolic
                max_time = Get time of maximum... onset offset Parabolic
                meanIntensity = Get mean... onset offset dB


                resultline$ =
"'label$"tab$"meanIntensity"tab$"min_int"tab$"min_time"tab$"max_int"tab$"max_time"tab$"'
                fileappend "'textfile$'" 'resultline$'
        endif
endfor

fileappend "'textfile$'" 'newline$'
endfor

# clean up

select all
Remove

#-------------
# This procedure finds the number of a tier that has a given label.

procedure GetTier name$ variable$
    numberOfTiers = Get number of tiers
    itier = 1
    repeat
         tier$ = Get tier name... itier
         itier = itier + 1
    until tier$ = name$ or itier > numberOfTiers
    if tier$ <> name$
         'variable$' = 0
    else
         'variable$' = itier - 1
    endif

        if 'variable$' = 0
                exit The tier called 'name$' is missing from the file 'soundname$'!
        endif

endproc
```

---

[5] Value changed to “FS” when measured intensity for the final syllable. Text output was renamed and saved in a separate directory.