University of Alberta

The effects of vocal loudness and speaking rate on voice-onset time in typically developing children and children with cochlear implants

by

Erica Lynn Knuttila

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Speech-Language Pathology

Department of Speech Pathology and Audiology

Spring 2011
Edmonton, Alberta

Abstract

This study explores the effects of manipulating vocal loudness and speech rate on voice onset time (VOT) in normal hearing children and two children with cochlear implants (CIs). 15 normal hearing participants and two participants with CIs produced all six stop consonants in the phrase “It’s a *C*od again” while speaking normally, softly, loudly, slowly, and quickly. Consonants were grouped into voiced and voiceless categories for comparison. Results indicated that the group of normal hearing children produced longer VOTs for voiceless stops than voiced across all conditions. When speaking loudly or quickly, VOT values were shorter than at normal levels. When speaking softly or slowly, VOT values were longer than at normal levels. The two children with CIs performed in a similar manner to the normal hearing group; however, VOTs produced by the six-year old participant were consistently longer than those of the normal hearing group across all conditions.

Acknowledgements

This thesis was funded by the Canadian Institute of Health Research.

I would first like to give special thanks to my co-supervisors, Drs. Carol Boliek and Melanie Campbell for all their support, guidance, and encouragement throughout this project. You ladies challenged me and taught me so much along the way.

Another thanks goes out to my examining committee members. Your advice and expertise really helped shape this project. Thank you to Dr. Bill Hodgetts for all your time and guidance especially with statistical analysis. I appreciate that you were always available to answer my questions. Thank you as well to Dr. Kathryn Ritter for all your help with participant recruitment. Your vast knowledge of children with cochlear implants was invaluable. Finally, thank you to Dr. Ming Zhang for the time you have spent on this project including the time spent with me at the Glenrose Rehabilitation Hospital reading patient files.

I would also like to thank a very special friend, Erin; the one person who understood exactly what I was going through. You were there through the ups and downs and always knew exactly what to say. I could not have made it through this experience without you.

Last, but certainly not least, I would like to send out an extra special thank you to my family and friends who have been there for me through it all. Your love and encouragement got me through the challenges and allowed me to succeed. I could not have asked for a more supportive team in my corner! I love you all.

Table of Contents

List of Tables

List of Figures

Introduction

Problem Statement

Speech intelligibility can be affected by a number of different factors, including nasality, articulation, vocal quality, and voice-onset time (VOT). VOT is “the interval between the release of an oral constriction (acoustically marked by a noise burst) and the onset of voicing (acoustically marked by the onset of periodic voicing)” (Whiteside, Dobbin, & Henry, 2003, p. 29). VOT is a temporal parameter of speech production, requiring precise coordination between the laryngeal and articulatory mechanisms. It is important for speech perception because it allows a listener to distinguish between the English voiced and voiceless stop consonants /p/ and /b/, /t/ and /d/, and /k/ and /g/. Without distinct VOTs for these consonants, words may be misconstrued and can potentially lead to unintelligible speech and subsequent communication breakdowns. Children born with sensorineural hearing loss, who gain access to auditory input through cochlear implantation, might have difficulty producing stop consonants with consistent VOT. There is limited research documenting the laryngeal-articulatory coordination required to produce consistent VOT distinctions in children with and without cochlear implants (CIs). The goal of the current study was to gain a better understanding of the stability, resilience, and precision of those laryngeal-articulatory gestures in children with normal hearing. The results were then used to help interpret the same in two children with CIs within the context of a case-study approach.

Development of VOT in Children with Normal Hearing

Five research studies have shown that children with normal hearing require a number of years to master adult VOT values; its stability occurs in the context of general maturation of motor speech abilities (Eguchi & Hirsh, 1969; Koenig, 2000; Whiteside et al., 2003; Whiteside & Marshall, 2001; Zlatin & Koenigsknecht, 1976). Taken together, these five studies have examined VOT in children from two to 13 years of age, but findings have differed. VOT is reported to have stabilized in boys and girls of seven to eight years (Eguchi & Hirsh, 1969); in girls nine to eleven years (Whiteside & Marshall, 2001); and in boys and girls eleven to 13 years (Whiteside et al., 2003). Given that the literature suggests that VOT stabilizes anywhere between seven years of age and puberty, it was reasoned that this study should be designed along a similar participant age span.

Development of VOT in Children with Compromised Hearing

Zlatin and Koenigsknecht (1976) suggested that stable VOT productions might be related to the perceptual skills necessary to discriminate between voiced and voiceless consonants. Without adequate perceptual skills, a child might be unaware of differences between cognate pairs of sounds.

Seven studies of VOT in children with CIs were found (Bharadwaj & Graves, 2008; Higgins, McCleary, Carney, & Schulte, 2003; Higgins, McCleary, & Schulte, 2001; Horga & Liker, 2006; Timgren, Boliek, & Campbell (in preparation); Tye-Murray, Spencer, & Gilbert-Bedia, 1995; Uchanski & Geers, 2003). The studies spanned ages six to 19 years. VOT in the bilabial position was examined in four of these studies; the /b/ and /p/ distinction was reported to be acquired by six to 19-year-olds in Tye-Murray et al. (1995), but acquired by only a minority (never more than half of the participants between six and 15 years of age) in the remaining studies (Higgins et al., 2003; Higgins et al., 2001; Timgren et al., in preparation). VOT in the alveolar position was examined in five of these studies; the /d/ VOT production was reported to be acquired by seven- to ten-year-olds by Timgren et al. (in preparation), but the remaining four studies showed that the /d/ and /t/ distinction was acquired by no more than half of the child participants (Bharadwaj & Graves, 2008; Uchanski & Geers, 2003) or not at all (Horga & Liker, 2006; Tye-Murray et al, 1995). None but Timgren et al. (in preparation) studied VOT for the velar position; none of their seven- to ten-year-old participants had acquired the /g/ and /k/ distinction.

In sum, the literature indicates that children with CIs vary in VOT acquisition, with only a minority of six- to nineteen-year-olds acquiring adult VOT values in the bilabial and alveolar positions. VOT in the velar position has been tested only once. This study was planned to observe VOT across all three places of articulation.

Neurological Mechanisms Underlying Speech Motor Control

Neuroimaging studies have shown that during speech, numerous areas of the brain are activated including the primary motor cortex of the mouth, lips, tongue and larynx (Huang, Carr, & Cao, 2002). Other cortical structures involved in speech production include the superior temporal gyrus, supramarginal gyrus, post-central gyrus (or primary sensory cortex), medial temporal gyrus, and superior parietal-temporal region (Chang, Kenney, Loucks, Poletto, & Ludlow, 2009). Subcortical structures also are involved in speech motor control and include the cerebellum, thalamus, putamen, caudate, periaqueductal grey, and lentiform nucleus (Chang et al., 2009; Jurgens, 2002).

Structures involved in the planning and sequencing of speech include the middle and inferior frontal gyri, cingulate cortex, insula, and supplementary motor area (Chang et al., 2009). It was found that reflexive central pattern generators in the periaqueductal grey and nucleus retroambiguus interact with higher controls in the central nervous system (e.g., primary motor cortex) to control laryngeal function during speech production (Ludlow, 2005; Zhang, Bandler, & Davis, 1995).

During speech production, a motor signal is sent from the primary motor cortex via corticobulbar tracts to the nucleus ambiguus where lower motor neurons activate the intrinsic muscles of the larynx in coordination with the respiratory system and articulators of the mouth (Ludlow, 2005; Simonyan & Jurgens, 2003). The integrity of this pathway is essential for the production of rapid speech acts such as voice onset (Ludlow, 2005).

As stated earlier, an essential component to VOT precision is adequate speech perception (Zlatin & Koenigskecht, 1976). Auditory and somatosensory input provides critical information for making online adjustments of the laryngeal, respiratory, and articulatory subsystems (Chang et al., 2009). Auditory and somatosensory feedback is sent back through the subcortical centers of the brain to the superior temporal gyrus (i.e., primary auditory cortex) and superior parietal-temporal region and inferior postcentral gyrus (i.e., primary somatosensory cortex) and inferior parietal cortex along the anterior supramarginal gyrus (Guenther, Ghosh, & Tourville, 2006). Moreover, Jurgens (2002) postulates that the periaqueductal grey may serve a gating function between audio and vocal information rather than a reflexive central pattern generating function because it receives direct sensory input from the superior and inferior colliculi, spinal trigeminal and solitary tract nuclei, and the dorsal horn of the spinal cord.

There are several opportunities for the speech system to make online adjustments for producing precise speech movements. There is some evidence to suggest that the feedback loops associated with motor adjustments range between 18 and 30 ms. More specifically, kinesthetic feedback from the laryngeal and articulatory muscles occurs within 18-25 ms of the speech act. This is followed by auditory feedback to the cortex within 20-30 ms after initial stimulation of the auditory system (Goffman, Ertmer, & Erdle, 2002; Larson, Altman, Liu, & Hain, 2008; Sharma, Nash, & Dorman, 2009).

Without auditory feedback, a child developing speech will not be able to integrate the kinesthetic feedback from the muscles of the larynx with the subtle differences (s)he hears when producing cognate pairs of plosives. This missing piece may be critical for mastery of VOT.

Children’s ability to produce accurate speech is dependent on the ability to receive auditory information. “Children learn to relate their own auditory outputs (auditory feedback) to their articulatory gestures, and they learn how the consequences of their articulatory gestures compare to sounds that are produced by other talkers” (Tye-Murray et al., 1995, p. 2460). Children with congenital deafness may have the opportunity to obtain some auditory feedback in the form of hearing aids or CIs. Tye-Murray et al. (1995) found that children with at least two years of CI experience were more likely to produce features such as nasality, voicing, and articulatory place correctly if they were able to perceive these features in a perceptual task. It is important to note that the subjects in that study had an average of 25% consonants correct in an audition-only perceptual condition, 59% consonants correct in an audition-plus-vision perceptual condition and 37% consonants correct in a speech production task. The authors suggested that CI experience “may have led to enhanced production of some features of articulation” but that more auditory experience may be required before these children begin to change their existing articulatory patterns (Tye-Murray et al., 1995, p. 2459). The results of this study will be framed within the context of the models of how auditory feedback influences speech production, as presented here.

Cochlear Implants and Their Function

While not the main focus, the present study also evaluated children with CIs. In order to understand this auditory input, it is important to understand how a CI device works. A CI is “an auditory prosthesis device for restoring hearing function…using electrical stimulation of [the] auditory nerve” (Kim, Kim, & Kim, 2007, p. 6352).

In order to encode speech, a signal is received through an external microphone typically worn behind the patient’s ear. This speech signal is then transmitted to a speech processor, which analyzes the signal and sends it to an external transmitter. The external transmitter is held in place via a magnet that connects it to a surgically implanted internal receiver. Radio frequency transmission sends the signal through the skin whereby the internal receiver decodes the signal and delivers electrical stimulation to the frequency-appropriate electrodes that are located in the cochlea. These electrodes are close to the auditory nerve, which sends the signal to the auditory centres of the brain (Chute & Nevins, 2000). CI devices used today are multi-channel devices (e.g., 12-22 channels), consisting of multiple electrodes positioned tonotopically in the cochlea (Zwolan, 2009). This allows the patient to receive information about pitch similar to someone with normal hearing. Sound frequency (or pitch) is necessary for understanding speech sounds. Currently, CI devices also use monopolar stimulation. Monopolar stimulation is achieved when electrical currents run through active and ground electrodes that are distant from each other. Particularly, the ground electrode is located externally to the cochlea whereas the active electrodes are located within the cochlea. In this type of stimulation lower level currents are required resulting in increased battery life and better thresholds in the speech processor.

The CI attempts to represent the damaged cochlea, activating the auditory nerve directly and providing the speaker with an auditory feedback loop. This newly created auditory loop can begin to provide auditory information so that the brain may integrate sound signals with kinesthetic feedback to adjust and master the precise timing necessary for the production of plosive consonants. One factor that may influence this precise timing feature is the speed at which the CI is able to process sound. Variables that can affect speed are: (a) intracochlear electrode array placement, (b) the capability of the CI sound processor, (c) age at implantation, and (d) speech processing strategies (Finley et al., 2008; Lohle et al., 1999; Santarelli et al., 2009; Zwolan, 2009). Behavioural outcomes such as word recognition, consonant discrimination, and speech perception have been used to test the fidelity of the implant. For example, Finley et al. (2008) suggested that word recognition scores for the CI population may improve through improvement of cochleostomy site selection and control of insertion depth during surgery. Lohle et al. (1999) found that in children implanted between the ages of two and 14 years, the youngest implanted group (age 2-4 years) achieved the best open-set speech perception results. Another study suggested that consonant discrimination ability in prelingually deaf adults wearing a unilateral CI was due to the implant and its ability to enhance cortical temporal processing (Roman, Canevet, Lorenzi, Triglia, & Liegeois-Chauvel, 2004). The results of this study will be framed within the context of the transmission speed characteristics of CIs, as described here.

Auditory Gap Detection

Gap detection testing is a method used to measure temporal processing (Michalewski, Starr, Nguyen, Kong, & Zeng, 2005). It is “the shortest time period over which the ear can discriminate two signals” (Chermak & Lee, 2005, p. 555). In normal hearing subjects, gap detection thresholds range from 3ms to 20ms, which is considered normal (Michalewski et al., 2005; Yalcinkaya, Muluk, Atas, & Keith, 2009). Specific norms for the Random Gap Detection Test (RGDT)

(Keith, 2000) show mean gap thresholds to be 6.0 to 7.8ms in normal hearing subjects (Chermak & Lee, 2005). In early-deafened subjects who have received CIs, research has shown that gap thresholds for this group ranged from 5ms or less to 95ms (Busby, Tong, & Clark, 1992). Given the large range, these researchers suggested that age at implantation and length of auditory deprivation might play a role in gap detection thresholds. However, Busby and Clark (1999) suggested the opposite; auditory deprivation and experience post-implantation may not play a role in gap detection thresholds. In relation to VOT and the perception of plosive consonants, it is believed that gap detection thresholds between 30ms and 40ms are adequate to encode and perceive this precise timing information (Busby & Clark, 1999). Based on the literature of gap detection reviewed above, this study was designed to measure the gap detection discrimination abilities of all participants.

Children with CIs have an altered experience in both perception and production of speech, which may not be as rich as that of normal hearing children. This population, depending on age at implantation, may experience neural development approximate to normal trajectories. On the other hand, if there is a significant delay in establishing an auditory feedback loop, children with CIs will have altered or more variable sensorimotor control systems than typically developing children. Moreover, it may be possible that the developmental trajectory for VOT is longer than that observed in normal hearing children. Speech production conditions can be manipulated in a variety of ways in an effort to gain a better understanding of laryngeal-articulatory control for VOT. The following section reviews studies that tested the stability, resilience, or precision of VOT laryngeal coordination using different speech tasks. These paradigms served to inform the selected speech tasks proposed in the present study.

Effects on VOT of Speech Condition Manipulations

Loudness. In adults with normal hearing, VOT values are stable, falling in the range of 0 ms to 25 ms for voiced tokens (not prevoiced) and 60 ms to 100 ms for voiceless tokens (Baken & Orlikoff, 2000; Lisker & Abramson, 1964). Studies have shown that manipulation of different components of speech production, such as speech breathing, lung volumes, and speaking rate can result in effects upon those VOT ranges. No studies were found that examined the effects of vocal loudness manipulations on VOT directly; therefore, a somewhat related study was used to generate predictions about loudness effects on VOT. Hoit, Solomon, and Hixon (1993) examined changes in lung volume across a single breath group and the effect it had on VOT in adults, a test of the resilience of the laryngeal coordination required to produce stable VOT values. The results of that study demonstrated that VOT tends to become longer at higher lung volumes and shorter at lower lung volumes. Vocal loudness may require higher lung volume initiations and higher subglottal pressures, therefore it was hypothesized that the manipulation of vocal loudness might also result in a similar effect to that observed when going to higher lung volumes. Specifically, it is hypothesized that an increase in vocal loudness will produce a longer VOT while a decrease in vocal loudness will produce shorter VOT values for any given plosive consonant.

Stathopoulos and Sapienza (1997) compared changes in the laryngeal and respiratory systems when adults and children altered their vocal loudness. The healthy subjects in that study were required to produce /pa/ repetitively. VOT measures were not taken, but glottal airflow, subglottal pressure, and sound pressure levels were collected using a Collins face mask and calibrated acoustic and aeromechanical system. These researchers found that children used more respiratory effort to produce vocal loudness, relying less on making changes within the laryngeal system to adjust glottal airflow or subglottal pressure. That may have been due to limited motor development of the laryngeal system and consequent difficulty making minor laryngeal adjustments for increasing loudness. Therefore, it was hypothesized that when children are asked to produce loud voices, they would have even less resilience of VOT laryngeal coordination than adults, and the effect would be even greater differences in VOT in children (compared to adults) between loud and normal conditions. Given that Stathopoulos and Sapienza (1997) succeeded in eliciting vocal loudness from their child participants, an *effort* task, it was hypothesized that the children in the current study would have no difficulty achieving a loud voice.

<u>Rate.</u> The effect on VOT of speaking rate, a test of precision of the laryngeal coordination required for VOT coordination, has been studied in adults (Volaitis & Miller, 1992). The researchers examined effects on syllable-initial plosive consonants across all places of articulation. They found that as a speaker decreased his or her rate of speech, syllable durations became longer. Consequently, VOT durations also increased for labial, alveolar, and velar places

of articulation. A larger range of VOT values were observed for plosive consonants when syllable duration increased. For slow speaking rates, it appears that the precision of the laryngeal coordination for VOT can be maintained to produce intelligible consonant distinctions with more timing latitude. Lane, Wozniak, and Perkell (1994) evaluated VOT across different syllable durations in four, postlingually deafened adults who used CIs. Though they did not manipulate speaking rate directly, these researchers found that the participants in their study naturally increased their speaking rate when asked to speak “louder”. It was found that as syllable durations decreased, VOT decreased. Other studies that have also evaluated VOT in adults at a variety of speaking rates have found that overall, when speaking rate was manipulated, voiceless tokens revealed greater changes in VOT than did voiced tokens (Kessinger & Blumstein, 1997; Lane et al., 1994; Miller, Green, & Reeves, 1986; Nagao & de Jong, 2007). Not surprisingly then, these researchers found that the voiced-voiceless categories shifted in the same direction as the change in mean VOT. In fast and slow rate conditions the laryngeal coordination precision for VOT is taxed. There is less latitude, that is, there is less time to produce an acceptable voiced plosive (0 to 25ms) than to produce an acceptable voiceless plosive (60 to 100 ms).

Again, given the greater time span in which voiceless plosives can be realized, it is not surprising that adults show wider variation in their production. Theodore, Miller, and DeSteno (2009) evaluated individual differences in adult VOT values for voiceless plosives when speaking rate was being manipulated. These researchers found that the magnitude of change in VOT varied between and within speakers across different speaking rates. No literature was found related to the effects of speaking rate on VOT laryngeal coordination in hearing children or in children with hearing loss.

Although not related to rate effects on VOT, Dwyer, Robb, Beirne, and Gilbert (2009) investigated the effect of an increased speaking rate on the nasality of prelingual, severely hearing-impaired speakers, aged twelve to nineteen years. All subjects were required to have used a hearing aid or CI at some point in their lives. The researchers found that by increasing speaking rate, hearing-impaired individuals showed improved nasality on subjective and objective measures of speech. However, the subjects in their study were unable to increase their speaking rate to that of a normal hearing control. The authors suggested that this might reflect the inability of some speakers with hearing loss to have sufficient speech motor control necessary to reach normal levels of precision. Nevertheless, taxing the speech production system resulted in an improvement in the resonance aspect of speech production. Is it reasonable to presume that a manipulation of rate will create the same results for VOT? It will be remembered that Stathopoulos and Sapienza (1997) found that children relied on their respiratory systems to make changes in vocal loudness whereas adults used a combination of the laryngeal and respiratory systems to do the same tasks. Taxing a child’s speech production system by having him speak with greater loudness, or more effort, resulted in less reliance on the laryngeal system, suggesting less resilience of the child’s laryngeal coordination. Conversely, taxing the child’s speech production system by having him speak at a faster rate, or with greater precision, resulted in an *improvement* in nasality. Are the two diametrically opposed? The current study intends to manipulate both *effort* (loudness) and *precision* (speaking rate) to observe their effects on VOT in children with normal hearing and in children with CIs. However, no assumptions are made on whether these manipulations will help or hinder the children's performance. Rather, results may serve to tease apart differences in two types of motor control parameters, the resilience and precision, of the coordination of the laryngeal-articulatory systems required to produce VOT distinctions. By understanding the type and degree of laryngeal-articulatory control exhibited by typical children, we can advance our understanding of underlying mechanisms associated with voice and speech movements observed in children with CIs.

Research Questions and Study Rationale

The purpose of this study was to evaluate the effects of manipulating speech rate and vocal loudness on laryngeal function in normal hearing children and children with CIs. This study aimed to better understand the control of voice and speech in typical children by providing insight into the effects of speech manipulations and production on a precise timing feature of speech. A second aim of this study was to better understand the association between the CI processing capabilities affecting the perception of the children with CIs and production of the precise timing feature of VOT. The main body of this study asked the following questions:

1. Do normal hearing children produce differences (comparable to adult values) in VOT between voiced and voiceless plosives when speaking normally?
2. Do normal hearing children produce differences in VOT between voiced and voiceless plosives when speaking with a soft voice or when speaking with a loud voice?
3. Do normal hearing children produce differences (comparable to adult values) in VOT between voiced and voiceless plosives when speaking slowly or when speaking quickly?
4. What effect does speaking with a soft voice have on VOT in voiced and voiceless plosives in normal hearing children? How does this compare to speaking with a loud voice? How does this compare to speaking slowly? To speaking quickly?
5. What effect does speaking with a loud voice have on VOT in voiced and voiceless plosives in normal hearing children? How does this compare to speaking slowly? To speaking quickly?
6. What effect does speaking slowly have on VOT in voiced and voiceless plosives in normal hearing children? How does this compare to speaking quickly?
7. What effect does speaking quickly have on VOT in voiced and voiceless plosives in normal hearing children?

In addition, a proof-of-concept study of two children with CIs was conducted, asking the same questions.

In the context of voicing features, it was predicted that normal hearing children will show longer VOT values when speaking louder or slower and will produce shorter VOT values when speaking softer or faster. It is difficult to predict how VOT will vary within groups across conditions with respect to voice features (e.g., voiced, voiceless). It is predicted that there will be interactions among voice features and condition. The results of two children with CIs were also compared descriptively to the results obtained for the group of normal hearing participants.

Research design

This study used a one-way within-subjects experimental design, with six dependent variables (VOT x 6 stop consonants), replicated across one group: Typical children (5-12 years). The independent variables were Condition having 5 levels: Normal, Soft voice, Loud voice, Slow rate, Fast rate. The dependent variables were VOT for 6 stop consonants (/b//p/, /d//t/, /g//k/) grouped into voiced (/b/, /d/, /g/) and voiceless (/p/, /t/, /k/) categories.

Methods

Participants

Fifteen typical children were sought as participants for the primary study based on a power calculation: $N = 16(15.62)/(118.4-101.2)^2 = 13$, for $\alpha = .05$, $\beta = 0.80$. They ranged in age from 5-12 years. Children had to pass a hearing screening test prior to participating. Criteria for passing included a positive response to three 20-dB HL tones at each of three frequencies: (1) 1000 Hz, (2) 2000 Hz, and (3) 4000 Hz.

Two children, ages six and 10 years, who used CIs were also included in the subsequent proof-of-concept study. Tables 1 and 2 present specifics related to each participant with CIs. Information includes date and age at implantation, make and model of the CI device, duration of CI use, most recent auditory thresholds and speech production assessment results for each child.

Table 1: *Time specifics of CI use in a 6-year old child and a 10-year old child. Gender, date of implantation, the make and model of CIs currently being used by each child, age at implantation, and duration of CI use are presented.*

			Age at Implantation		**Duration of CI use**	
Subject Code	**Gender**	**Make and model of implant**	**Right ear**	**Left ear**	**Right ear**	**Left ear**
M0601	Male	**Make:** Advanced Bionics **Model:** HR90K b/l **Processor:** Harmony **Speech processing strategy:** HiRes-P with Fidelity 120; wears an FM system	1;6	4;6	5;5	2;6
M1001	Male	**Make:** Advanced Bionics **Model:** CII CI (R) HR90K CI (L) **Processor:** Prefers Platinum series **Speech processing strategy:** HiRes Fidelity 120 b/l; uses Phonak EasyLing transmitter with receivers coupled to both PSP devices (FM system)	1;1	7;9	9;2	3;5

Table 2: *Performance results in a 6-year old child and a 10-year old child with CIs. Age when tested in the current study, most recent auditory thresholds, and most recent speech production assessment results are presented.*

Subject Code	**Age when tested**	**Gender**	**Most recent auditory thresholds for both ears**	**Most recent speech production assessment report**
M0601	6	Male	In soundfield, responded to narrow band noise stimuli from 230-4000Hz btwn 15 and 20dBHL; responses for R were similar or slightly poorer than B/L; L was 5-10DB louder than B/L; Speech perception @ 60DBSPL: B/L: PBK words: 80%; PBK phomemes: 95%; HINT-C quiet: 98%; HINT-C +10S/N: 86%; HINT-C +5S/N: 75%	WNL as of June 2010; with occasional speech sound errors; able to discriminate Ling 6 sounds in B/L condition and with either CI alone; able to discriminate between min pairs of words in a closed set with over 90% accuracy; the only errors were on final consonant ID in either place or voicing
M1001	10	Male	Soundfield testing: responded to narrow band noise from 250-4000Hz between 20-25dBHL in B/L and CI1 alone; responses in CI2 fell btwn 30-40dBHL; Speech perception: CI1/bilateral: PBK words: 72%/76%; PBK phonemes: 86%/89%; HINT-recorded +7 dB S/N: 84%/88%; WIPI: 56% and HINT-C sentences in quiet: improved from 31% at 60dBSPL to 54% at 70dBSPL	N/A

These children were oral communicators and attended school in their local communities. They were both clients of the Glenrose Rehabilitation Hospital in

Edmonton, which has the responsibility for monitoring all children with CIs in Central and Northern Alberta. An oral communicator is a person who is able to communicate with others using spoken language in addition to gestures and other non-verbal behaviors. Children with CIs who had concomitant diagnoses such as cognitive delay or Autism Spectrum Disorder were not considered for the current study.

This study was approved by the Health Research Ethics Board at the University of Alberta in Edmonton, Alberta, Canada. All participants gave written (aged 7 years and up) or verbal (aged 5-6 years) assent. The participants' parents also provided written consent for participation.

Equipment

Equipment and materials used for measurement included: (a) a unidirectional condenser microphone (audio-technica, model AT8537), (b) a digital audio recorder capable of sampling at 44 kHz (Tascam Digital Audio Tape Recorder, model DA-P1), (c) a digital sound level meter for measuring loudness levels (RadioShack®, model 33-2055, and (d) a portable audiometer for screening the hearing of the typical subjects (Maico, model MA-25).

Procedures

Stimulus materials

The carrier phrase used in the perception and production tasks was "It's a *C*od again." This phrase was used for all six English plosives in the normal, soft voice, loud voice, slow rate, and fast rate conditions. The carrier phrase chosen for this study parallels the phrase used in a previous study by Lane et al. (1994) which evaluated VOT across different syllable durations in four adults who used CIs. The current study replicated the procedures used in the Lane et al. (1994) study for manipulation of speaking rate.

Data collection

In order to perform the necessary measurements, the following steps were followed. The 15 typical participants began with a hearing screening. They were asked to wear a set of headphones and listen for very soft beeping sounds, first in one ear and then the other. They were asked to raise one hand anytime they heard a beep.

Perceptual testing was conducted during the same test session in order to demonstrate that the children were able to hear what they were being expected to produce. Children listened to a high quality digital recording of an adult female speaker saying all of the speech tokens in the carrier phrase described above. Each phrase was produced once for all of the loudness and rate conditions used in the experiment. All participants listened to the recording with headphones (normal hearing participants) or in a free-field, sound booth (participants with CIs). After each sentence, children were asked to point to the target word (all options provided on a response form with a corresponding picture) they thought they heard and the researcher recorded each response on a separate form. Each response form contained the six plosive consonants. Each consonant was judged in each of the five speaking conditions for a total of 30 judgments. These responses were used to describe VOT perception in percent accuracy by task, for each participant.

Further perceptual testing also was conducted using the Random Gap Detection Test (RGDT) (Keith, 2000). All participants completed the test with headphones (normal hearing participants) or in a free-field, sound booth (participants with CIs). Participants were required to listen to tones at 500Hz, 1000Hz, 2000Hz, and 4000Hz with 10 different gaps at intervals between 0 and 40ms in random order. Next, participants were required to listen to white noise clicks with 10 different gaps at random intervals between 0 and 40ms. After each tone or click, participants were asked to state whether they heard one or two tones and/or hold up their fingers to indicate one or two tones. Responses were recorded on the response form for the RGDT (Keith, 2000).

Following perceptual testing, a microphone was mounted on the participant's forehead and secured with tape and a headband to ensure comfort and a fixed mouth-to-microphone distance. A digital sound pressure meter was placed 30 cm from the participant's mouth. A research assistant recorded dB SPL values during each sentence production. One measurement was taken in the middle of each sentence (as close to "*C*od" as possible) to represent the peak dB SPL for that sentence. Sound pressure level was not a dependent variable but was used to verify that participants realized actual differences in dB SPL among normal, loud, and soft conditions. All acoustic recordings were sampled at 44 kHz.

Participants were asked to repeat the carrier phrase: "It's a /*C*/od again." five times for each of the following six consonants: /p/, /b/, /t/, /d/, /k/, and /g/, for a total of 30 productions, at a normal rate and loudness. The sentence was modeled each time by the researcher in a repetition-type paradigm. The participants were then asked to repeat the carrier phrase speaking twice as loud as normal (loud voice), followed by speaking half as loud as normal (soft voice). Lastly, the participants were asked to repeat the carrier phrase speaking twice as fast as normal (fast rate), followed by speaking half as fast as normal (slow rate). For the purpose of maintaining consistency in administration across participants the same researcher demonstrated speaking normal, speaking twice as loud as normal and half as loud as normal, as well as speaking twice as fast as normal and half as fast as normal. The researcher provided the participants with CIs with letter and picture prompts for each trial to maximize the accuracy of the target word produced. The normal loudness/rate condition was presented first for all participants and then just one production of each of the six plosives was elicited between each loudness and/or rate condition in order to "reset" the participant to his/her normal loudness/rate before moving into the next condition. The order of the experimental conditions was randomized across participants. The order of consonant production was randomized within conditions to control for potential practice and order effects.

Data analysis

Sound pressure levels (dB SPL) recorded for each sentence were used to confirm changes in vocal loudness for soft and loud conditions. The protocol used by Dwyer et al. (2009) was employed in the current study to confirm changes in rate of speech for slow and fast conditions. This protocol used TF32 acoustic analysis software (Milenkovic, University of Wisconsin, Madison, WI) where each sentence was displayed on a computer monitor. A vertical cursor was placed at the onset of voicing of the first syllable in the sentence, and a second cursor was placed at the offset of voicing of the last syllable in the sentence. This time interval was recorded as the total sentence duration. The total number of syllables in the sentence was divided by the total sentence duration to obtain the speaking rate (number of syllables per second). This protocol was used to verify the speaking rate conditions of each participant.

Each perceptual response form was analyzed for overall percent consonants correct (including all conditions) and used to describe the perceptual accuracy for each participant. Each RGDT (Keith, 2000) was analyzed for either the shortest or the mean shortest gap (ms) for clicks and tones, respectively. Mean shortest gap (ms) for tones was calculated by averaging the shortest gap (ms) across all frequencies (500, 1000, 2000, and 4000Hz). The shortest gap (ms) for clicks was determined as the shortest gap detected in the 10 different gaps that were presented from 0ms to 40ms. Mean loudness levels (dB SPL) were calculated using the values recorded from the digital sound pressure meter for each condition and subject. Mean rates (syllables/second) were calculated based on each rate-manipulated condition and subject.

All audio recordings were edited prior to acoustic analysis. Acoustic samples were digitized at 44 kHz and analyzed using *Praat* acoustic analysis software (Boersma & Weenink, 2010). VOTs were determined for all trials for each consonant token produced in each condition. One cursor was placed at the onset of the consonant burst, and a second cursor was placed at the onset of voicing for the vowel. VOT was calculated for the time between the two cursors. Figure I.1 in Appendix I shows the measurement. Individual means and standard deviations were calculated for all VOTs (ms) by token for each condition and subject. These means were then used to calculate voiced and voiceless VOT (ms) means and standard deviations for each condition and subject. Basic descriptive statistics for VOTs (ms) were calculated and used to guide the selection of appropriate parametric statistical tests for within-group comparisons (voiced vs. voiceless for each condition). The data met the criteria for parametric analysis, therefore, a 2-voicing by 5-condition repeated measures ANOVA with 25 *a priori* paired comparisons was employed.

Intra-rater and inter-rater reliability

Intra-rater reliability was obtained via the reanalysis of VOTs for ten percent of the data.

Inter-rater reliability was obtained via the reanalysis of VOTs for ten percent of the entire data sample by a second researcher. Before any data analysis began, the two researchers discussed the measurement criteria and procedures for VOT measurements.

Results

Normal Hearing Participants

This study used a one-way within-subjects design to evaluate the effects of vocal loudness and speech rate on voiced vs. voiceless categories of VOT (ms) in a group of normal hearing children (5-12 years). A 2-voicing (voiced, voiceless) x 5-condition (normal, soft voice, loud voice, slow rate, fast rate) repeated measures ANOVA was employed to evaluate these effects.

Table 3 shows the descriptive data presented individually for each normal hearing participant. The percent correct indicated for the perceptual task was derived from each participant's ability to identify the speech tokens (e.g., bod, pod, dod, tod, god, kod) used in the study. The mean shortest and shortest gap detections, for tones and clicks respectively, are reported in ms. The average dB SPL and syllable/sec are presented and indicate that each participant successfully performed the speech manipulations required in the production tasks. Additionally, the group mean for dB SPL values for each condition are presented. It was determined that in the soft condition participants decreased their loudness on average by one and a half times their normal loudness level and in the loud condition, participants increased their loudness on average by two times. Statistical analysis using paired t-tests were run to compare loudness levels across loudness and rate conditions (Bonferroni correction of p-value < 0.0125 was used). The soft and loud conditions showed dB SPL values that were significantly different from normal (soft: $t(14) = 6.20$, $p < 0.001$ (1-tailed); loud: $t(14) = -17.60$, $p < 0.001$ (1-tailed)), the slow rate condition was on the border of being significantly different (slow: *t* (14) = -2.88, *p* =0.012 (2-tailed), and the fast rate condition did not show significant dB SPL changes from normal (fast: *t* (14) = -1.05, *p* =0.313 (2-tailed)).

Table 3: *Descriptive data for each participant. Scores for the speech perceptual task are presented in percentage correct responses. The average shortest gap detected on tones and the shortest gap detected for clicks are presented for each participant. The average shortest gap detected for tones for subjects F0501, F0601, M1103 included only 2/4 trials and subject F0902 included only 1/4 trials, due to unreliable responses. Average dB SPL for each loudness condition and average syllables/sec for each rate condition also are presented. NR indicates that the shortest gap could not be recorded due to unreliable responses (e.g., when a child said s/he heard two tones for the 0ms gap, one tone for the 2ms and 5ms gaps and then said two tones for the 10ms gap. The researcher was unable to gauge the true gap detection threshold in such cases.)*

			Perceptual Task	**Random Gap Detection Test**		**Production Tasks - Mean dB SPL per condition**					**Production Tasks - Mean Rate (syllables/sec) per condition**		
Subject Code	**Gender**	**Age (yrs)**	**% Consonants Correct**	**Mean Shortest Gap (msec)* - Tones**	**Shortest Gap (msec) - Clicks**	**Normal**	**Soft**	**Loud**	**Slow**	**Fast**	**Normal**	**Slow**	**Fast**
F0501	Female	5	57%	7.5	NR	<50.0	<50.0	56.8	<50.0	<50.0	5.0	2.5	7.3
F0502	Female	5	83%	NR	NR	59.4	51.6	76.4	59.0	58.9	5.3	2.3	8.1
F0601	Female	6	90%	5	2	<50.0	<50.0	64.8	<50.0	<50.0	5.1	2.2	6.4
F0801	Female	8	73%	2	10	52.9	<50.0	67.6	51.6	53.6	4.7	2.9	6.9
F0802	Female	8	97%	2.75	5	53.3	<50.0	70.1	59.9	53.7	5.0	1.9	9.2
M0801	Male	8	97%	2.75	5	54.7	52.6	72.5	56.7	55.4	4.8	2.0	6.2
M0802	Male	8	97%	6.75	5	51.3	51.2	65.3	54.5	51.6	6.0	2.2	9.0
F0901	Female	9	93%	NR	NR	52.5	<50.0	66.2	56.2	52.1	5.3	1.8	7.9
F0902	Female	9	100%	8	15	54.8	51.3	72.5	65.0	60.9	6.2	2.1	8.9
F1001	Female	10	100%	2.75	5	55.1	<50.0	66.7	54.6	55.2	5.5	2.6	8.0
M1001	Male	10	100%	2	5	54.3	<50.0	74.1	61.2	57.5	5.6	1.9	8.6
M1101	Male	11	90%	3.5	5	56.4	50.9	66.6	54.6	56.2	5.7	2.4	8.0
M1102	Male	11	100%	2	5	54.3	<50.0	73.0	58.2	52.1	4.9	2.1	7.9
M1103	Male	11	97%	5	2	56.9	53.0	72.6	61.0	57.1	5.5	2.3	7.9
F1201	Female	12	100%	2	NR	55.3	51.2	73.6	57.0	54.6	4.8	2.2	7.2

*Includes shortest gap for 500Hz, 1000Hz, 2000Hz and 4000Hz tones

Group mean dB SPL	53.9	50.3	69.3	56.5	54.5

A Pearson correlation on the initial and re-analysis of data for intra-rater reliability resulted in a value of $r = 0.94$, $p < 0.01$. A Pearson correlation on the two analyses for inter-rater reliability resulted in a value of $r = 0.94$, $p < 0.01$. Table 4 provides group means and standard deviations derived for VOT for each condition. In the normal condition, the normal hearing group produced a mean VOT of 16.44 ms for the voiced tokens which fell into the 0-25ms range of adult norms for voiced tokens. A mean VOT of 72.53 ms for the voiceless tokens in the normal condition also fell within the 60-100 ms range of adult norms for voiceless tokens (Baken & Orlikoff, 2000; Lisker & Abramson, 1964). For the loudness and rate conditions the normal hearing group produced mean VOT values that fell within the adult range (0-25ms for voiced tokens and 60-100ms for voiceless tokens) (Baken & Orlikoff, 2000; Lisker & Abramson, 1964) except for in the *voiceless token, loud voice* condition and *voiceless token, fast rate* condition. Group means were compared to group medians (Appendix J) for each condition. For all conditions except the *voiceless token, soft voice* and *voiceless token, loud voice* conditions, the mean and median were within one and three ms of each other. Distributions for each condition were evaluated to ensure that the appropriate statistical tests were employed. Box and whisker plots, skewness, kurtosis and standard error values are presented in Appendix K, Figure K.1 and Tables K.1, K.2a, K.2b, and K.2c, respectively. The distribution of VOT values derived from *voiced tokens, normal rate* and *voiced tokens, fast rate* were classified as leptokurtic. In addition, the distribution of VOT values derived from *voiced tokens, slow rate* and *voiced tokens, fast rate* were classified as positively and negatively skewed, respectively. Parametric tests were applied to the data with confidence of their appropriateness.

Table 4: *Group means and standard deviations for VOT measurements for all conditions*

Condition	**Mean (S.D.)**
Voiced tokens, normal rate and loudness	16.438 (8.137)
Voiced tokens, soft voice	29.704 (9.973)
Voiced tokens, loud voice	10.225 (3.987)
Voiced tokens, slow rate	21.315 (10.408)
Voiced tokens, fast rate	13.953 (10.605)
Voiceless tokens, normal rate and loudness	72.532 (19.029)
Voiceless tokens, soft voice	82.114 (21.709)
Voiceless tokens, loud voice	45.909 (10.491)
Voiceless tokens, slow rate	105.080 (22.503)
Voiceless tokens, fast rate	40.837 (10.751)

Figure 1 shows group error bar plots for VOT for each condition. These plots reveal more variability with voiceless tokens than voiced but show no 95% confidence interval overlap between voiced and voiceless tokens across comparable conditions (i.e., *voiced tokens, soft voice* vs. *voiceless tokens, soft voice*). However, these plots also reveal 95% confidence interval overlap for different conditions within voicing categories.

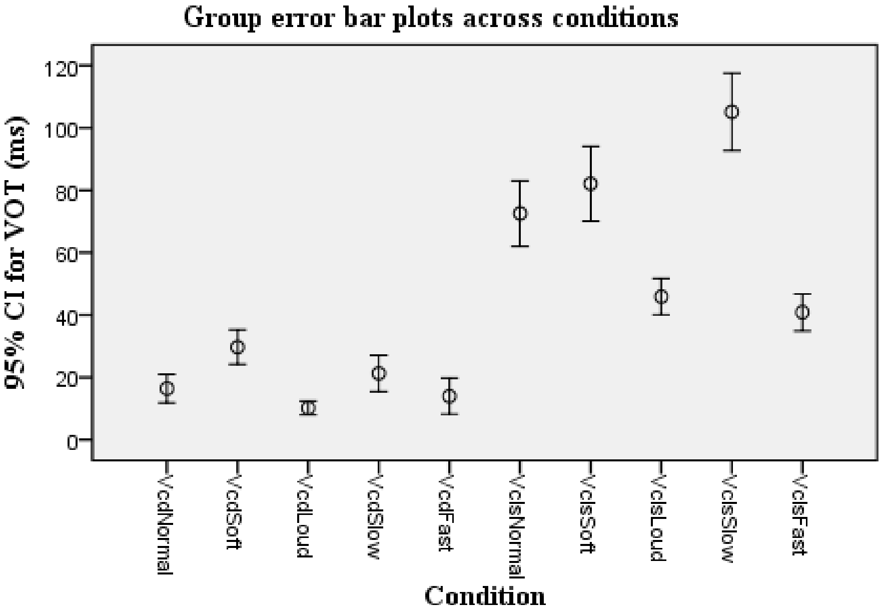

Figure 1: *Group error bar plots across conditions. VcdNormal = voiced tokens, normal condition, VcdSoft = voiced tokens, soft voice condition, VcdLoud = voiced tokens, loud voice condition, VcdSlow = voiced tokens, slow rate condition, VcdFast = voiced tokens, fast rate condition, VclsNormal = voiceless tokens, normal condition, VclsSoft = voiceless tokens, soft voice condition, VclsLoud = voiceless tokens, loud voice condition, VclsSlow = voiceless tokens, slow rate condition, and VclsFast = voiceless tokens, fast rate condition.*

A within-group, repeated measures ANOVA indicated a non-significant Mauchley's test of sphericity ($p = 0.59$) for the full statistical model indicating sphericity of the data. Main effects were found for voicing, $F(1, 14) = 357.34$, $p < 0.001$, $\eta^2 = 0.96$ and condition, $F(4, 56) = 76.21$, $p < 0.001$, $\eta^2 = 0.85$. A significant interaction was found for voicing x condition, $F(4, 56) = 30.20$, $p < 0.001$,. $\eta^2 = 0.68$. Twenty-five *a priori* paired *t*-tests were conducted using a Bonferroni correction such that $p \leq 0.002$ was needed to reach statistical significance. Table 5 presents all of the *t*-test comparisons, confidence intervals, *p* values and cohen's *d* results for each comparison. As can be seen in Table 5, seven of the 24 comparisons did not reach significance.

Table 5: t-*tests for the 25* a priori *comparisons. Mean differences, upper bound 95% confidence interval, lower bound 95% confidence interval,* t-*value,* p-*value, and Cohen's* d *are presented. Bolded comparisons indicate statistical significance. Highlighted comparisons indicate a large effect size with a trend towards significance.*

	Paired Differences					
		95% Confidence Interval of the Difference				
Pair	Mean difference (msec)	Upper	Lower	*t*	*p*	*d*
VcdNormal - VclsNormal	**-56.09**	**-47.51**	**-64.68**	**-14.01**	**.000**	**-4.13**
VcdSoft - VclsSoft	**-52.41**	**-40.69**	**-64.13**	**-9.59**	**.000**	**-3.31**
VcdLoud - VclsLoud	**-35.68**	**-30.57**	**-40.80**	**-14.96**	**.000**	**-4.93**
VcdSlow - VclsSlow	**-83.76**	**-71.03**	**-96.50**	**-14.11**	**.000**	**-5.09**
VcdFast - VclsFast	**-26.82**	**-19.06**	**-34.59**	**-7.41**	**.000**	**-2.53**
VcdSoft - VcdNormal	**13.27**	**16.71**	**9.83**	**8.27**	**.000**	**1.47**
VcdSoft - VcdLoud	**19.48**	**24.77**	**14.19**	**7.90**	**.000**	**2.79**
VcdSoft - VcdSlow	8.39	13.67	3.11	3.41	.004	0.82
VcdSoft - VcdFast	**15.69**	**21.34**	**10.04**	**5.96**	**.000**	**1.54**
VclsSoft - VclsNormal	9.58	18.33	.84	2.35	.034	0.47
VclsSoft - VclsLoud	**36.21**	**48.54**	**23.87**	**6.29**	**.000**	**2.25**
VclsSoft - VclsSlow	**-22.97**	**-13.61**	**-32.32**	**-5.26**	**.000**	**-1.04**
VclsSoft - VclsFast	**41.28**	**52.77**	**29.78**	**7.70**	**.000**	**2.54**
VcdLoud - VcdNormal	**-6.21**	**-2.85**	**-9.58**	**-3.96**	**.001**	**-1.02**
VcdLoud - VcdSlow	**-11.09**	**-5.89**	**-16.29**	**-4.58**	**.000**	**-1.54**
VcdLoud - VcdFast	-3.79	.83	-8.41	-1.76	.101	-0.53
VclsLoud - VclsNormal	**-26.62**	**-14.79**	**-38.45**	**-4.83**	**.000**	**-1.80**
VclsLoud - VclsSlow	**-59.17**	**-47.92**	**-70.42**	**-11.28**	**.000**	**-3.59**
VclsLoud - VclsFast	5.07	13.09	-2.95	1.36	.196	0.48
VcdSlow - VcdNormal	4.88	9.12	.63	2.46	.027	0.53
VcdSlow - VcdFast	7.30	12.82	1.78	2.84	.013	0.70
VclsSlow - VclsNormal	**32.55**	**42.09**	**23.01**	**7.32**	**.000**	**1.57**
VclsSlow - VclsFast	**64.24**	**76.83**	**51.66**	**10.95**	**.000**	**3.86**
VcdFast - VcdNormal	-2.42	1.96	-6.81	-1.19	.255	-0.26
VclsFast - VclsNormal	**-31.69**	**-21.56**	**-41.83**	**-6.71**	**.000**	**-2.13**

However, four of these comparisons could be considered having a statistical trend based on *p* value in combination with its power (*d*). The comparisons not reaching significance included: (a) *voiceless token, soft voice condition* and *voiceless token, normal condition*, (b) *voiceless token, loud voice* condition and *voiceless token, fast rate* condition, and (c) *voiced token, fast rate* condition and *voiced token, normal* condition. Possible trends were found for: (a) *voiced token,* *soft voice* condition and *voiced token, slow rate* condition, (b) *voiced token, loud voice* condition and *voiced token, fast condition*, (c) *voiced token, slow rate* condition and *voiced token, normal condition*, and (d) *voiced token, slow rate* condition and *voiced token, fast rate* condition.

Table 6 presents the 25 *a priori* comparisons in relation to the original predictions.

Table 6: *Group comparisons of VOT by voicing and condition. The specific comparison is listed in the first column. The direction of the prediction is indicated in the second column by a less than (<) or (>), indicating the direction of VOTs with less than = shorter and greater than symbols = longer durations. << or >> equals a change reaching statistical significance. No predictions were made when VOTs in a comparison were expected to change in the same direction, relative to duration. Comparisons shaded in gray went in the opposite direction of the prediction.*

Comparison	**Prediction**	**Finding**
VcdNormal - VclsNormal	VcdNormal < VclsNormal	VcdNormal < < VclsNormal
VcdSoft - VclsSoft	VcdSoft < VclsSoft	VcdSoft < < VclsSoft
VcdLoud - VclsLoud	VcdLoud < VclsLoud	VcdLoud < < VclsLoud
VcdSlow - VclsSlow	VcdSlow < VclsSlow	VcdSlow < < VclsSlow
VcdFast - VclsFast	VcdFast < VclsFast	VcdFast < < VclsFast
VcdSoft - VcdNormal	VcdSoft < VcdNormal	VcdSoft > > VcdNormal
VcdSoft - VcdLoud	VcdSoft < VcdLoud	VcdSoft > > VcdLoud
VcdSoft - VcdSlow	VcdSoft < VcdSlow	VcdSoft = VcdSlow
VcdSoft - VcdFast	No prediction	VcdSoft > > VcdFast
VclsSoft - VclsNormal	VclsSoft < VclsNormal	VclsSoft = VclsNormal
VclsSoft - VclsLoud	VclsSoft < VclsLoud	VclsSoft > > VclsLoud
VclsSoft - VclsSlow	VclsSoft <VclsSlow	VclsSoft < < VclsSlow
VclsSoft - VclsFast	No prediction	VclsSoft > > VclsFast
VcdLoud - VcdNormal	VcdLoud > VcdNormal	VcdLoud < < VcdNormal
VcdLoud - VcdSlow	No prediction	VcdLoud < < VcdSlow
VcdLoud - VcdFast	VcdLoud > VcdFast	VcdLoud < VcdFast
VclsLoud - VclsNormal	VclsLoud > VclsNormal	VclsLoud < < VclsNormal
VclsLoud - VclsSlow	No prediction	VclsLoud < < VclsSlow
VclsLoud - VclsFast	VclsLoud > VclsFast	VclsLoud = VclsFast
VcdSlow - VcdNormal	VcdSlow > VcdNormal	VcdSlow > VcdNormal
VcdSlow - VcdFast	VcdSlow > VcdFast	VcdSlow > VcdFast
VclsSlow - VclsNormal	VclsSlow > VclsNormal	VclsSlow > > VclsNormal
VclsSlow - VclsFast	VclsSlow > VclsFast	VclsSlow > > VclsFast
VcdFast - VcdNormal	VcdFast < VcdNormal	VcdFast = VcdNormal
VclsFast - VclsNormal	VclsFast < VclsNormal	VclsFast < < VclsNormal

The results revealed 11 comparisons that met the prediction. 10 comparisons did not meet what was predicted. Of those 10 comparisons, 6 comparisons went in the opposite direction of what was predicted and 4 comparisons showed no difference between the conditions. There were 4 comparisons that could not be predicted *a priori*. These were comparisons where the conditions were expected to move in the same direction from normal (i.e., it was predicted that voiceless token, soft voice condition and voiceless token, fast rate condition would produce VOT values that were shorter than normal). Since it was not known which condition would have the greatest effect on VOT, no prediction could be made about which would result in the longest or shortest VOT.

Participants with Cochlear Implants

Table 7 shows the descriptive data derived for each participant with CIs. The perceptual data provide information related to each participant's ability to discriminate the speech tokens (e.g., bod, pod, dod, tod, god, cod) used in the study. The loudness and rate data indicate that each participant successfully performed the speech manipulations required in the production tasks.

Table 7: *Descriptive data for each participant with CIs. Scores for the speech perceptual task are presented in percentage correct responses. The average shortest gap detected on tones and the shortest gap detected with clicks are presented for each participant. The average shortest gap detected for tones for subject M0601 included only 3/4 trials due to an unreliable response. Average dB SPL for each condition and average rate (syllables/sec) for the rate tasks also are presented. NR indicated unreliable responses.*

			Perceptual Task	Random Gap Detection Test		Production Tasks - Mean dB SPL per condition					Production Tasks - Mean Rate (syllables/sec) per condition		
Subject Code	Gender	Age (yrs)	% Consonants Correct	Mean Shortest Gap (msec)* - Tones	Shortest Gap (msec) - Clicks	Normal	Soft	Loud	Slow	Fast	Normal	Slow	Fast
M0601	Male	6	100%	14.0	5.0	60.6	54.6	68.8	57.3	60.9	4.2	1.7	6.7
M1001	Male	10	93%	17.5	NR	54.2	50.6	69.8	55.2	53.2	5.5	2.3	7.3

*Includes shortest gap for 500Hz, 1000Hz, 2000Hz and 4000Hz tones

Table 8 provides mean VOT (ms), standard deviations and median VOT (ms) for each condition produced by the two participants with CIs. If the reader compares Table 8 to Table 4, it can be seen that both participants with CIs produced VOT patterns similar to the normal hearing group across conditions: (a) average VOTs for voiced conditions were shorter than the average VOTs for voiceless conditions, (b) soft voice conditions (voiced and voiceless) produced VOT values that were longer than VOTs observed in the normal conditions (voiced and voiceless), (c) loud voice conditions (voiced and voiceless) produced VOT values that were shorter than normal conditions (voiced and voiceless), (d) slow rate conditions (voiced and voiceless) resulted in increased VOT values compared to normal conditions (voiced and voiceless), and (e) fast rate conditions (voiced and voiceless) resulted in decreased VOT values compared to normal conditions (voiced and voiceless). However, some of the values for VOT were well outside the normal hearing

Table 8: *Individual means, standard deviations and medians for all conditions for participants with CIs.*

	CI Subject			
	M0601		M1001	
Condition	Mean VOT (ms) (S.D.)	Median VOT (ms)	Mean VOT (ms) (S.D.)	Median VOT (ms)
Voiced token, normal	21.84 (9.75)	22.12	13.64 (5.34)	11.59
Voiced token, soft voice	37.23 (4.22)	39.16	29.15 (5.29)	31.82
Voiced token, loud voice	16.04 (6.52)	16.00	7.93 (3.07)	8.08
Voiced token, slow rate	28.26 (10.42)	29.66	20.82 (7.84)	17.19
Voiced token, fast rate	26.91 (9.05)	28.90	7.73 (18.91)	15.54
Voiceless token, normal	102.42 (20.18)	95.31	64.88 (24.28)	64.22
Voiceless token, soft voice	95.19 (31.41)	99.52	76.23 (21.08)	86.43
Voiceless token, loud voice	76.43 (15.08)	82.40	36.31 (9.41)	40.85
Voiceless token, slow rate	219.87 (50.01)	217.02	113.32 (24.41)	111.54
Voiceless token, fast rate	66.87 (13.59)	67.52	36.79 (8.66)	38.69

group. To evaluate this more closely, each participant will be compared individually to the normal hearing group.

6-year old male with cochlear implants

Figure 2 presents the mean VOT (ms) for each condition in a 6-year old child with CIs in comparison with the mean VOT (ms) for each condition in the normal hearing group.

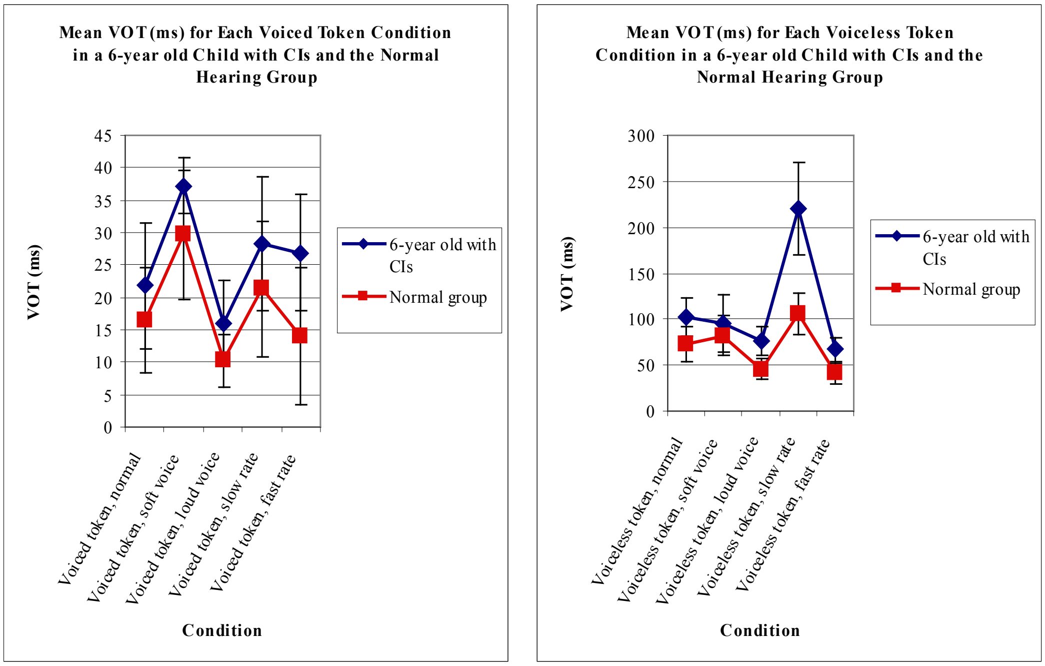

Figure 2: *Mean VOT (ms) for each condition in a 6-year old child with cochlear implants compared to the normal hearing group. Error bars represent one standard deviation above and below the means for the participant with CIs and the normal hearing group.*

Compared to the normal hearing group, a 6-year old participant with CIs produced VOT values that were generally longer in every condition. As can be seen in Figure 2, however, the 6-year old child with CIs was within one standard deviation of the normal group mean for the following tasks: (a) *voiced token,* *normal* condition, (b) *voiced token, soft voice* condition, (c) *voiced token, slow rate* condition, and (d) *voiceless token, soft voice* condition.

10-year old male with cochlear implants

Figure 3 presents the mean VOT (ms) for each condition in a 10-year old child with CIs. This child produced VOT values that were shorter or equal to those values of the normal hearing group in each condition except the *voiceless token, slow rate* condition which was longer than the normal hearing group mean. As can be seen in Figure 3, the 10-year old child with CIs was within one standard deviation of the normal group mean for VOT in every condition.

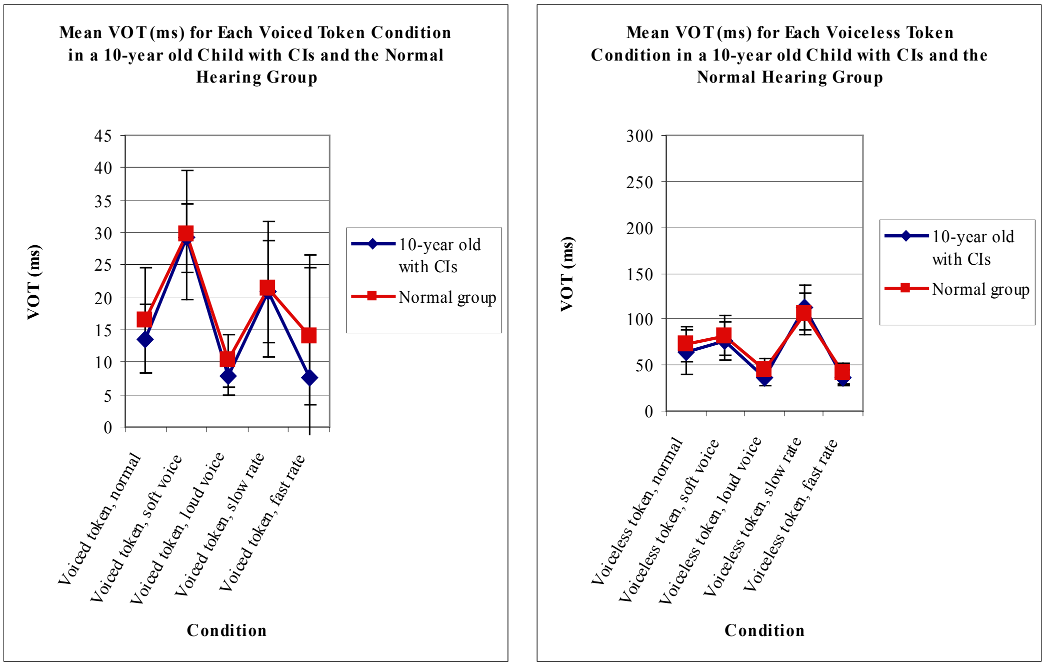

Figure 3*: Mean VOT (ms) for each condition in a 10-year old child with CIs compared to the normal hearing group. Error bars represent one standard deviation above and below the means for the participant with CIs and the normal hearing group.*

Discussion

This study examined VOT (ms) in a group of 15 normal hearing participants and two participants with CIs when vocal loudness and speech rate were manipulated. The results of this study provided insight into the resilience and precision of how normal hearing children coordinate laryngeal-articulatory gestures under various speaking conditions and established an experimental paradigm that was used as proof-of-concept that the coordination of laryngeal-articulatory gestures in children with CIs could be studied in the same way.

The major finding from the current study was that normal hearing children demonstrated resilient and precise coordination of the laryngeal-articulatory system that enabled them to compensate for changes in vocal loudness (*effort*) and speaking rate (*precision*), indicating refinement of motor coordination in these dimensions. In addition, the two children with CIs exhibited the ability to compensate for vocal loudness and speech rate manipulation resulting in VOT patterns that were similar to those observed in the hearing children. More specific interpretations of results are presented in the context of group, voicing, and condition.

Normal Hearing Participants

Voiced vs. voiceless tokens

In general, production of the voiceless tokens revealed more VOT variability than the voiced tokens. According to Kewley-Port and Preston (1974), production of voiceless stops requires more precise control of timing between the separately innervated oral and laryngeal articulators and are, therefore, more difficult to produce than voiced stops. Moreover, these researchers suggested that this precise timing is not necessary when producing voiced stops because "adduction of the vocal folds can be achieved any time during [stop] closure…and oscillation will still begin only upon [stop] release." (p.140). In addition to more precise control of oroarticulatory and laryngeal systems, there is evidence suggesting that voiceless stops require more muscular activation of intrinsic and extrinsic laryngeal muscles than needed for voiced stops. Hirose and Gay (1972) conducted EMG studies on the inter-arytenoid muscles and found that adduction of the vocal folds for the production of voiceless stops requires more force and less time for movement than for voiced stops. In addition, the posterior cricoarytenoid muscle is activated for a longer period of time during the production of voiceless stops (Hirose & Gay, 1972). It is possible that the variability in VOT observed for voiceless tokens in the current study may be due to a lack of motor refinement for precise timing as suggested by Kewley-Port and Preston (1974) as well as the muscle activation dynamics associated with voiceless token movements (Hirose & Gay, 1972). Other studies have reported similar variability in VOT for voiceless relative to voiced plosives in children (Eguchi & Hirsh, 1969; Koenig, 2000, 2001; Uchanki & Geers, 2003; Whiteside & Marshall, 2001). VOT stability has been found to occur as young as seven years of age (Eguchi & Hirsh, 1969); however, Koenig (2000) suggested that VOT in voiceless stops may not stabilize until puberty. Adult VOT norms for voiced (without prevoicing) and voiceless tokens in the normal condition are 0ms to 25ms and 60ms to100ms, respectively (Baken & Orlikoff, 2000; Lisker & Abramson, 1964). The results of the normal conditions in the current study fell into these ranges which indicated that, overall, the normal hearing participants have reached adult-like values when speaking normally.

Loudness conditions

It was hypothesized that normal hearing participants, when speaking with a loud voice, would produce longer VOT values whereas, when speaking with a soft voice, would produce shorter VOT values. The results of the current study were in the opposite direction of this prediction, but based on dBSPL versus lung volume events (Hoit et al. 1993). Specifically, it was found that increased vocal loudness resulted in VOT values that were shorter than normal and decreased vocal loudness resulted in VOT values that were longer than normal. Woo (1996) evaluated normal vocal fold behavior in adults who produced sustained phonation on /ee/ at modal pitch, high pitch, low pitch, and loud phonation using videostrobolaryngoscopy. He found that when phonating at a louder than normal level, adult vocal folds opened and closed faster than when phonating normally, thus producing an increased rate of glottal area opening and closing and a faster adduction of the folds. Additionally, he found that in the loud condition, the vocal folds were in the closed position for a longer duration within a cycle. These results may have been in part due to changes in subglottal pressure needed to increase vocal loudness. Increasing vocal loudness results from increases in subglottal pressure. An increase in subglottal pressure will cause the vocal folds to increase their muscular length- tension and subsequently move to midline (towards a closed position) to prevent them from being "blown" open (Holmberg, Hillman, & Perkell, 1988; Ludlow, 2005). The fact that the vocal folds are already approximating midline, in combination with the evidence showing faster vocal fold movement during loud phonations, could result in the shorter VOTs observed during loud productions in the current study. Though it was not evaluated specifically by Woo (1996), it can be inferred that when speaking in a soft voice, the vocal folds would open and close at a slower rate than when phonating normally and maintain an open position for a relatively longer period of time. An increase in the open phase of the glottal cycles could be the result of a decrease in subglottal pressure and, subsequently, lower vocal fold tension associated with soft phonations (Baken & Orlikoff, 2000; Holmberg et al., 1988; Ludlow, 2005). Lower muscular tension on the vocal folds results in less pull on each fold towards midline. It follows that under these circumstances, more time is required for vocal folds to reach midline resulting in longer VOT for soft productions as was observed in the current study.

Another important potential factor to consider is laryngeal airflow which is affected by subglottal pressure and airway resistance. In adult women it has been found that as vocal loudness increased, glottal resistance increased, though laryngeal airflow changed inconsistently (Baken & Orlikoff, 2000). It is important to note that the data derived from the current study cannot be interpreted in the context of direct measures of lung volume events (e.g., lung volume initiations, terminations or excursions), prevailing subglottal pressures, laryngeal airflow or oral airflow. However, these aerodynamic factors likely contributed to the VOT values obtained in the current study (Lofqvist, 1992).

Voice onset times were more variable when produced with a soft voice regardless of voicing type. Once again prevailing subglottal pressures, vocal fold length tension characteristics, and airflow factors may contribute to laryngeal-articulatory control needed for achieving loudness targets (Baken & Orlikoff, 2000; Holmberg et al., 1988; Lofqvist, 1992; Ludlow, 2005; Kewley-Port & Preston, 1974). The variable VOT values obtained from the soft loudness condition in the current study indicated that precise coordination of laryngeal-articulatory control has not yet become stereotypical in this age range. When asked to speak half as loud as normal, children typically spoke in a whisper. A prompt to "keep your voice on" often was given, even to older children. Perhaps the control needed to maintain voicing while speaking quietly required greater control than voicing while speaking loudly. Fisher and Swank (1997) evaluated phonation threshold pressure and suggested that, based on previous research demonstrating variability in pressure peaks of phonation at soft loudness levels, "it is possible that precise phonation near a physiological limit (such as phonation threshold) is difficult even for rehearsed speakers and singers" (p. 1122). It follows that the developmental trajectory for laryngeal-articulatory control when speaking softly may be longer than the trajectory for speaking at a louder than normal level and would affect VOT values.

Manipulating vocal loudness had a similar effect on VOTs for voiced and voiceless productions as evidenced by similar timing patterns. For example, VOTs were shorter for voiced and voiceless tokens spoken at louder than normal levels whereas VOTs were longer for voiced and voiceless tokens spoken at softer than normal levels. However, voiceless tokens spoken at softer than normal loudness levels did not reach statistical significance when compared to VOTs produced at normal loudness levels. Taken together, these comparisons suggest that children are capable of adjusting the timing features of laryngeal-articulatory movements when producing speech differing in vocal loudness. It appears that laryngeal-articulatory coordination is challenged most when children are asked to produce voiceless tokens at a softer than normal voice level perhaps because of the more complex movements associated with voiceless plosives and soft speech (Fisher & Swank, 1997; Kewley-Port & Preston, 1974).

It is difficult to know if the VOT values associated with loudness manipulation approximate those of adults. Stathopoulos and Sapienza (1997) reported measures of fundamental frequency and glottal airflow from children and adults when producing syllable trains at different loudness levels. The authors found that children tended to use the respiratory system (e.g., go to higher lung volumes) when asked to produce syllables at louder than normal levels. They suggested that the respiratory adjustment was larger than the laryngeal adjustment for achieving speaking targets varying in vocal loudness. Adults in their study made more laryngeal than respiratory adjustments to vocal loudness manipulation. Children in the current study did make laryngeal-articulatory adjustments to varying loudness targets. Without respiratory data, it is difficult to discern the relative amount of adjustment between respiratory-laryngeal subsystems.

Rate conditions

The overall results of VOTs derived from slow and fast speaking rate conditions for the voiced and voiceless tokens were as predicted. That is, when speaking at a fast rate, VOT values were shorter and when speaking at a slow rate, VOT values were longer than those obtained at normal speaking rates. These results support previous research on VOTs produced by adults under similar conditions (Kessinger & Blumstein, 1997; Miller et al., 1986; Nagao & de Jong, 2007; Volaitis & Miller, 1992). Mean VOT values for /b/ and /p/ at a fast speaking rate in adults has been found to be 13ms and 63ms, with a range of 0-39ms and 20-119ms, respectively. Mean VOT values for /b/ and /p/ at a slow speaking rate in adults has been found to be 15ms and 95ms, with a range of 0-39ms and 40-149ms, respectively (Kessinger & Blumstein, 1997). Though the results of the current study include all three places of articulation within the voiced and voiceless categories, it was found that the mean VOT for rate conditions fit within the ranges provided for /b/ and /p/ in adults. This indicated that the normal hearing children in the present study have adult-like VOT productions even when speaking rate is manipulated.

Another finding was that speaking rate did not have as big an effect on VOTs derived from voiced tokens as it did on voiceless tokens. Both fast and slow speaking rates significantly changed VOTs derived from voiceless tokens. The findings in the current study are supported by data from previous studies also showing that changes in speaking rate have a greater effect on voiceless than on voiced tokens (Kessinger & Blumstein, 1997; Miller et al., 1986; Nagao & de

Jong, 2007; Volaitis & Miller, 1992). Kessinger and Blumstein suggested that voiced tokens exist within a small VOT range and therefore, have "little acoustic space in which exemplars may vary" (p. 162). In contrast, these authors stated that voiceless tokens could be produced within a much larger VOT range and subsequently, have more acoustic space within which to move (e.g., more timing and distance degrees of freedom). These same researchers also describe rate effects on VOT in the context of articulatory gestures. They stated that there are "intrinsic limitations on the articulatory gestures for producing stop consonants…and these cannot be affected by changes in speaking rate" (Kessinger & Blumstein, 1997, p. 163). Specifically, in order to produce significant changes in VOT (i.e., pre-voiced stop for voiced plosives) when one increases speaking rates, a speaker would be required to alter the timing between the burst and the onset of voicing. In contrast, for slow rate conditions where VOT generally gets longer, producing a significant change would require one to delay the onset of voicing. This would involve the precise ability to maintain the vocal folds in an abducted state which would likely cause changes to the internal structure of the voiced stop (e.g., aspiration noise that is generally characteristic of voiceless stops).

Although the effects for voiced tokens in the rate conditions were not as large as those in the voiceless conditions it should be noted that VOTs derived from voiced tokens spoken at slower and faster rates exhibited a trend for longer and shorter durations, respectively. However, VOTs derived from the *voiced token, fast rate* condition were not significantly shorter than those from the normal condition nor did they appear to follow a trend towards significance. Upon evaluating the raw individual data, it was found that 5 out of the 15 normal hearing participants produced inconsistent negative values for VOT in the fast condition. These negative values were due to the absence of an apparent burst. In these participants, the lack of burst at the faster speaking rate may be due to decreases in muscle movement amplitude sometimes found when speaking rate is increased (Guenther, 1995).

Across loudness and rate conditions

Although specific hypotheses were not stated for the interactions between conditions and voicing, it was predicted that interactions would exist. The results showed a number of across condition comparisons that reached statistical significance (see Table 6). Overall, it appeared that manipulating vocal loudness had a greater effect on VOT than did speaking rate, for voiced plosives. As explained previously, there may be inherent limitations of rate effects on VOTs for voiced plosives (Kessinger & Blumstein, 1997; Miller et al., 1986; Nagao & de Jong, 2007; Volaitis & Miller, 1992). Changing vocal loudness in the production of voiced tokens does not involve making changes to the articulatory gestures (e.g., maintaining the vocal folds in an abducted state or producing aspiration noise) but rather changes in the effort that drives the speech act. In effect, these changes in effort produce greater changes to VOT.

On the other hand, it appeared that manipulating speaking rate had a greater effect on VOT than did changing vocal loudness, for voiceless plosives. Stated previously, voiceless tokens can be produced within a much larger VOT range and subsequently, have more acoustic space within which to move or more degrees of freedom in timing of laryngeal-articulatory movements (Kessinger & Blumstein, 1997). Because the vocal folds are positioned farther away from midline for voiceless plosives and have a relatively larger range of motion following the burst, it is reasoned that timing perturbations (speaking rate) would have a larger impact on VOT than would loudness perturbations. Therefore, physical position and mechanical attributes of the vocal folds associated with the production of voiceless plosives may account for more significant speech rate effects on VOT.

Another possible explanation comes from Wohlert & Hammen (2000) who found that changes in speaking rate were accompanied by changes in vocal loudness; however, changes in vocal loudness were unaccompanied by changes in speaking rate. Thus it is reasoned that when manipulating speaking rate, VOT is affected by both changes in rate and loudness. These combined effects may account for the significant speech rate effects on VOT in voiceless plosives..In the current study, the dB SPL calculations and t-test comparisons revealed that vocal loudness changed significantly in the slow rate but not in the fast rate condition. However, due to the limitations of the sound level meter (i.e., only being able to detect loudness values above 50 dB SPL) a true loudness level for the normal condition was unattainable. As a result, no definite conclusions can be drawn about whether the vocal loudness in the slow rate condition had an effect. Taken together, the data from the current study indicate that loudness and rate perturbations have differential effects on VOTs from voiced and voiceless tokens.

Perception versus production

As can be seen in Table 3, not all of the participants achieved 100% accuracy in the perceptual task. It appeared that the younger children in the group tended to have less accurate perceptual skills than the older children; however, when looking at the box and whisker plots (Figure K.1) it appeared that not one child was consistently different from the others. Looking specifically at the two children with the lowest perceptual accuracy (F0501 and F0801), the following was found: (a) F0501 had misperceptions of tokens in all conditions except the fast condition and was an outlier for VOTs produced in the *voiced token, normal* condition and the *voiced token, slow condition,* and (b) F0801 had misperceptions of tokens in all conditions except the normal condition but was never an outlier in the production tasks. A thorough analysis of error patterns between the perceptual results and the production (VOT) results was not conducted but should be considered in future research.

Another result to note is that in the *voiceless token, loud voice* condition and *voiceless token, fast rate* condition mean VOT (ms) fell between the voiced and voiceless boundaries. Thus these tokens may not easily be distinguished as either a voiced or voiceless token. As a result, words may be misconstrued. It would have been interesting to explore if these conditions in particular caused ambiguity in the perception of adult listeners.

A 6-year old Male with Cochlear Implants

Auditory perception

This participant achieved 100% accuracy in the perceptual task and normal gap detection thresholds (i.e., less than 20ms). Additionally, recent speech production assessment results indicate that other than in final consonant position, this child has no difficulty producing age-appropriate speech sounds. This descriptive data revealed that this participant was able to adequately hear the stop consonants he was being asked to produce in the speech production tasks. As well, it was implied from these results that this participant receives adequate auditory input from his CIs on a daily basis. It is possible that VOT values falling outside the range of the hearing group in the current study were the result of motor control refinement issues as opposed to CI processing capabilities. Unfortunately, it was not possible to evaluate this participant's auditory feedback when producing and adjusting his own speech but based on the results, it appeared that this child was still in the process of refining the laryngeal-articulatory coordination needed to produce voiceless consonants. This made it even more difficult for him when his articulatory and laryngeal systems were pushed outside the operating range.

VOT manipulations

This participant, while demonstrating similar patterns of VOT change when manipulating speech rate and vocal loudness, consistently produced VOT values that were longer than the normal hearing group. This pattern was especially true for rate conditions and voiceless tokens. Macken and Barton (1979) evaluated VOT longitudinally as a measure of the acquisition of the voicing contrast in four normal hearing children (age 1 year at the time of the first session and age 2 at the time of the second session). These researchers found that acquisition of the voicing contrast is divided into three stages, whereby Stage III was marked by adult-like values for voiceless stops. It was also found that some children had a tendency to exceed adult-like VOT values for a period of time prior to reaching adult-like values. This research can shed some light on the 6-year old participant with CIs.

It is likely that the 6-year old participant with CIs in the current study was in a phase like some of the children in the study by Macken and Barton (1979). Data from this child indicated that VOT values were generally longer than those produced by his typically developing counterparts; however, these differences were not uniform across conditions (see Figure 2). It is important to consider here that this participant had only five years of hearing experience in his right ear and two years hearing experience in his left ear at the time of testing and therefore, had much less hearing experience than most of the children in this study. It is reasonable to assume that this lack of hearing experience is related to the differences revealed between this 6-year old participant and the normal hearing group.

When comparing this 6-year old participant with CIs to his peers, matched for hearing experience (two 5-year old female participants), the results revealed that generally these three participants produced equal VOT values in the voiced conditions, but in the voiceless conditions, the 6-year old participant with CIs generally produced VOT values that were longer than the 5-year old participants. However, all three participants with 5-years hearing experience exhibited high inter and intra-subject variability. Zlatin and Koenigskecht (1976) found that the range of VOT values for /t/ for 6-year old children was 60 to 110ms and for /d/ was 0 to 20ms. Whereas the current study grouped all the voiced and voiceless tokens into two categories, the raw data for this participant with CIs revealed a mean VOT of 95.31ms for /t/ in the normal condition and 22.12ms for /d/ in the normal condition. Overall, this child appeared to produce VOTs for alveolar stops within the same range when compared to age-matched typical children. It is likely that the young children in the current study, including the 6-year old participant with CIs, continue to refine and master the increased muscular involvement and precision necessary to produce consistent VOT in voiceless tokens regardless of condition (Kewley-Port & Preston, 1974). The differences found between this participant and his peers matched for hearing experience are likely due to the lack of auditory feedback received this participant’s first year of life. Despite being matched for hearing experience it appeared that the child with CIs had more difficulty with articulatory-laryngeal coordination. In a study by Tye-Murray et al. (1995) children with CIs who had more than two years of CI experience were more likely to reveal correct voicing and articulatory place when they could perceive these features accurately. These researchers proposed that despite this, more auditory experience was necessary for these children to produce adult-like articulatory patterns. It would be interesting to know if, given more hearing experience, the 6-year old participant with CIs in the current study would reach adult-like productions.

A 10-year old Male with Cochlear Implants

Auditory perception

It is interesting that this participant scored 93% on the perceptual task, had longer gap detection thresholds which fell within normal, and had higher auditory thresholds than the 6-year old participant with CIs yet his VOT values were more closely in line with those of the normal hearing group. The researcher concludes that this is likely due to duration of auditory experience. This participant has had more time to use auditory feedback to monitor his speech and adjust and develop his motor speech skills. This time has afforded him the opportunity to reach the levels of skill needed to manipulate vocal loudness and speech rate as well as normal hearing children, despite having only one CI for most of his life. Again, it appears that the speech processing capabilities (i.e. processing speeds) of this participant's CIs do not inhibit him from achieving normal levels of speech, even precise timing features such as VOT. Further interpretation of the speech production results are discussed now.

VOT manipulations

In comparison to the normal hearing group, the 10-year old participant with CIs produced similar trends in VOT changes; VOT increased from normal in the soft voice and slow rate conditions and decreased from normal in the loud voice and fast rate conditions. In comparison to the 6-year old participant with

CIs, this 10-year old participant with CIs produced VOT values that were consistently lower but within one standard deviation of the mean VOT values in the normal hearing group. This participant had nine years of auditory experience in the right ear and three years in the left ear at the time of testing. Whereas the left ear had less auditory experience than any of the normal hearing children in the study, recent chart review revealed that this child does better with his right ear and relies on it more in everyday life. This participant with CIs produced VOT values that were approximately equal to those of the two 9-year old normal hearing females who participated in the current study. This generally occurred for both voiced and voiceless tokens. Bharadwaj & Graves (2008) conducted a study on a group of prelingually deaf children with CIs and used the following values from Uchanski and Geers (2003) for comparison to normal hearing children: the range of VOT values for /t/ used for children with 7-11 years of CI use was 22 to 145ms and for /d/ was -45 to 38ms. The raw data for alveolar stops for the 10-year old participant with CIs in the current study revealed a mean VOT of 64.22ms for /t/ in the normal condition and 11.59ms for /d/ in the normal condition. Both of these values fall within what is considered the normal range for this place of articulation in normal hearing children. Based on these findings, it appeared that the developmental trajectory for VOT in this participant with CIs is on track with that observed in normal hearing children.

One final comment relates to age at implantation. Both participants with CIs received their first auditory experience at the age of one year. This has likely also played a major factor in the success both have achieved in speech production and perception as it is known that the earlier a child receives auditory experience, the greater his/her chances are of reaching normal limits in speech production. This is confirmed in the numerous studies conducted on VOT in children with CIs (Bharadwaj & Graves, 2008; Higgins et al., 2003; Higgins et al., 2001; Horga and Liker, 2006). The results of previous studies revealed that the activation of hearing through a CI can improve VOT, but that this improvement has been variable within and across speakers. Importantly, none of the children tested in these studies were implanted as young as the participants in the current study demonstrating that age at implantation may in fact, be a crucial component to success. As a result of this discrepancy, it is difficult to compare the present findings to findings from any these studies. Zlatin and Koenigsknecht (1976) suggested that adequate perceptual skills are necessary to discriminate and produce stable differences between voiced and voiceless consonants which confirms the current results.

Limitations and Future Research

Although this study aimed to be as consistent as possible with the collection and analysis of data, some factors may have affected the results. Throughout the data collection process, a number of assistants were utilized to record dB SPL values from the sound pressure meter. All assistants were instructed to record the value that appeared during the “*C*od” portion of the carrier phrase but given the likelihood of inter-rater differences, the reliability of this cannot be guaranteed. Fortunately, dB SPL values were used merely to demonstrate that the children manipulated vocal loudness as expected. A related issue was that loudness values were not controlled. That is, the children manipulated vocal loudness based on the researcher's model rather than being expected to reach a certain loudness level (dB SPL) in each loudness condition. Whereas all of the participants manipulated their loudness, the magnitude of these changes was not uniform throughout. The dB SPL ranged from 56.8 to 76.4 dB SPL in the loud condition, from <50.0 to 53.0 dB SPL in the soft condition, and from <50.0 to 56.9 dB SPL in the normal condition. This corresponded to an increase of over twice the normal loudness for tokens produced in the loud condition and a decrease of one and a half times the normal loudness in the soft condition. However, caution should be used when comparing these differences in magnitude as the true dB SPL for the soft condition was not measurable due to the limitations of the sound level meter. That is, the sound level meter could not record speech that was less than 50 dB SPL. The participants often went below this level when speaking softly. If VOT was affected by changes in loudness, how can one be certain that the magnitude of this change was not affected by the magnitude of the changes in loudness? On the other hand, the current study aimed to determine if and how children were able to manipulate speech on their own and was not concerned so much with the magnitude of change.

Another factor that may have affected the results is that the current study did not measure respiration, subglottal pressure, laryngeal airflow, and oral airflow during the speaking tasks. These data would have added to a more detailed understanding of the aeromechanical environment associated with rate and loudness perturbations. Additionally, the current study did not control for pitch, particularly in the loudness conditions. Fundamental frequency (f0) measures for vowel productions produced for each condition would have revealed whether or not a significant shift in f0 was related to changes in VOT. Several of the aeromechanical measurements are difficult to acquire during running speech so speech tasks would need to be modified to meet the assumptions of aeromechanical assessment. Nevertheless, information derived from these additional variables would have added depth to the interpretation of the present findings and should be considered for future research in this area.

A limitation of the current study already previously mentioned is that any stop consonant that did not appear to have a burst was considered prevoiced and therefore received a negative VOT value. The tokens that did not contain a burst should have been removed from the data to provide consistency in analysis. It is possible that these values affected group means for the fast rate conditions in the voiced category.

The current study comprised a sample of 15 children of which 9 were female and 6 were male. Whiteside and Marshall (2001) evaluated VOT in children aged 7, 9, and 11 years and found that between the ages of 9 and 11, children revealed sex-linked differences in VOT. Whereas the current study had age- and gender-matched subjects for ages 8 and 10, the remaining age levels included did not (i.e. 5-6 years, 9 years, and 11-12 years). This presents a limitation in the number of ways the data could have been evaluated. Although the focus of this study was not on sex-linked differences, having produced a sample conducive to such would have expanded the insight into the way boys versus girls coordinate laryngeal-articulatory gestures when manipulating speech rate and vocal loudness. In the future, research should consider the differences between boys and girls when conducting studies to evaluate manipulation of speech production and the effect this has on different aspects of speech such as VOT.

One last limitation of this study was the limited number of participants with CIs recruited. With a sample size of two, it was not possible to get a thorough idea of how children with CIs coordinate laryngeal-articulatory gestures (VOT) when manipulating vocal loudness and speech rate relative to normal hearing children. However, the results from the two children with CIs provided a proof–of-concept related to the sensitivity of the experimental protocol to detect small changes in VOT when manipulated by loudness and rate conditions. Obviously, future research should focus on a larger group of children with CIs to further our knowledge about how this population coordinates laryngeal-articulatory gestures (VOT) when manipulating speech rate and vocal loudness in the context of (a) VOT perception, (b) processing capabilities of cochlear prosthetics as determined by device specifications and patient interface, and (c) age at implantation and duration of hearing experience.

Future research should also include analysis of error patterns between children's perception and production skills. This will provide further insight into the role of auditory feedback on speech motor control and development. Additionally, future research should have adults judge children's productions of stop consonants when vocal loudness and speaking rate are manipulated to determine if changes in VOT affect others' perceptions. For example, the *voiceless token, loud voice* condition and *voiceless token, fast rate* condition produced mean VOT (ms) values that fell between the two ranges for voiced and voiceless tokens (between 25ms and 60ms). It would have been interesting to see how others perceived these productions.

Conclusion

Children, like adults, are able to manipulate their vocal loudness and speech rate to produce changes in VOT. Speaking in a loud voice or at a fast rate decreases VOT while speaking in a soft voice or at a slow rate increases VOT. However, the magnitude of these changes is not uniform across conditions or even within conditions, across voicing. Children appear to have a longer developmental trajectory in refining voiceless tokens and this is likely a result of the increased effort required to do so. Additionally, when altering speaking rate, it appears that this manipulation requires very precise control which can be difficult even for adults.

The two participants with CIs in the current study provided proof-of-concept that this particular experimental protocol is capable of measuring small changes in VOT due to loudness and rate manipulations in high functioning children with CIs. The protocol was able to identify increases in VOT when speaking at a slow rate or with a soft voice and decreases in VOT when speaking at a fast rate or with a loud voice. It appeared that age at implantation and duration of CI use/auditory experience were the two major factors involved in the success of speech production and perception in these two participants. Further research needs to be conducted to increase this understanding of how children with CIs coordinate laryngeal-articulatory gestures (VOT) when manipulating speech rate and vocal loudness. The results should be interpreted in the context of (a) age at implantation, (b) duration of CI use/auditory experience, (c) CI processing capabilities, (d) auditory thresholds, and (e) VOT perception. In doing so, conclusions may be drawn which may inform voice and speech therapies used with this pediatric population.

Bibliography

Baken, R., & Orlikoff, R. (2000). *Clinical measurement of speech and voice* (2nd ed.) San Diego, CA: Singular.

Bharadwaj, S. V. & Graves, A. G. (2008). Efficacy of the discreteness of voicing category (DOVC) measure for characterizing voicing errors in children with cochlear implants: A report. *Journal of Speech, Language, and Hearing Research*, 51, 629-635.

Boersma, P. & Weenink, D (2010). Praat: doing phonetics by computer [Computer program]. Version 5.2.16, retrieved 01 May 2010 from http://www.praat.org/

Busby, P. A., & Clark, G. M. (1999). Gap detection by early-deafened cochlear-implant subjects. *Journal of Acoustical Society of America*, 105 (3), 1841-1852.

Busby, P. A., Tong, Y. C., & Clark, G. M. (1992). Psychophysical studies using multiple-electrode cochlear implant patients who were deafened early in life. *Audiology*, 31, 95-111.

Chang, S., Kenney, M. K., Loucks, M. J., Poletto, C. J., & Ludlow, C. L. (2009). Common neural substrates support speech and non-speech vocal tract gestures. *Neuroimage*, 47, 314-325.

Chermak, G. D., & Lee, J. (2005). Comparison of children’s performance on four tests of temporal resolution. *Journal of the American Academy of Audiology*, 16, 554-563.

Chute, P. M., & Nevins, M. E. (2000). Cochlear implants in children. In M. Valente, H. Hosford-Dunn, & R. J. Roeser, *Audiology Treatment* (pp. 511-535). New York: Thieme.

Dwyer, C. H., Robb, M. P., O'Beirne, G. A., & Gilbert, H. R. (2009). The influence of speaking rate on nasality in the speech of hearing-impaired individuals. *Journal of Speech, Language, and Hearing Research*, 52, 1321-1333.

Eguchi, S and Hirsh, I. J. (1969). Development of speech sounds in children. *Acta Oto-laryngologica supplemental*, 257, 5-51.

Finley, C. C., Holden, T. A., Holden, L. K., Whiting, B. R., Chole, R. A., Neely, G. J., et al. (2008). Role of electrode placement as a contributor to variability in cochlear implant outcomes. *Otology and Neurotology*, 29, 920-928.

Fisher, K. V. & Swank, P. R. (1997). Estimating phonation threshold pressure. *Journal of Speech, Language, and Hearing Research*, 40(5), 1122-1129.

Goffman, L., Ertmer, D. J., & Erdle, C. (2002). Changes in speech production in a child with a cochlear implant: Acoustic and kinematic evidence. *Journal of Speech, Language, and Hearing Research, 45*, 891-901.

Guenther, F. H. (1995). Speech sound acquisition, co-articulation, and rate effects in a neural network model of speech production. *Psychological Review*, 102(3), 594-621.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, 96, 280-301.

Higgins, M. B., McCleary, E. A., Carney, A. E., & Schulte, L. (2003). Longitudinal changes in children's speech and voice physiology after cochlear implantation. *Ear and Hearing, 24*(1), 48-70.

Higgins, M. B., McCleary, E. A., & Schulte, L. (2001). Articulatory changes with short-term deactivation of the cochlear implants of two prelingually deafened children. *Ear and Hearing, 22*(1), 29-45.

Hirose, H., & Gay, T. (1972). The activity of the intrinsic laryngeal muscles in voicing control. *Phonetica*, 25 (3), 140-164.

Hoit, J. D., Solomon, N. P., & Hixon, T. J. (1993). Effect of lung volume on voice onset time (VOT). *Journal of Speech and Hearing Research, 36*, 516-521.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *Journal of Acoustical Society of America*, 84 (2), 511-529.

Horga, D., & Liker, M. (2006). Voice and pronunciation of cochlear implant speakers. *Clinical Linguistics & Phonetics, 20*(2), 211-217.

Huang, J., Carr, T. H., & Cao, Y. (2001). Comparing cortical activations for silent and overt speech using event-related fMRI. *Human Brain Mapping, 15*, 39-53.

Jurgens, U. (2002). Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews, 26*, 235-258.

Keith, R. W. (2000). Random Gap Detection Test. St. Louis, Missouri: Auditec Inc.

Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics*, 25, 143-168

Kewley-Port, D., & Preston, M. (1974). Early apical stop production: A voice onset time analysis. *Journal of Phonetics*, 2, 195-210.

Kim, K. H., Kim, J. H., & Kim, D. H. (2007). An improved speech processor for cochlear implant based on active nonlinear model of biological cochlea. *Engineering in Medicine and Biology Society, 29th Annual International Conference of the IEEE EMBS,* Cite Internationale, Lyon, France. 6352-6355.

Koenig, L. L. (2000). Laryngeal factors in voiceless consonant production in men, women, and 5-year-olds. *Journal of Speech, Language, and Hearing Research,* 43, 1211-1228.

Koenig, L. L. (2001). Distributional characteristics of VOT in children's voiceless aspirated stops and interpretation of developmental trends. *Journal of Speech, Language, and Hearing Research*, 44, 1058-1068.

Lane, H., Wozniak, J., & Perkell, J. (1994). Changes in voice-onset time in speakers with cochlear implants. *Journal of Acoustical Society of America, 96*(1), 56-63.

Larson, C. R., Altman, K. W., Liu, H., & Hain, T. C. (2008). Interactions between auditory and somatosensory feedback for voice control. *Exp Brain Res., 187*, 613-621.

Lisker, L. & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.

Lofqvist, A. (1992). Acoustic and aerodynamic effects of interarticulator timing in voiceless consonants. *Language and Speech*, 35 (1, 2), 15-28.

Lohle, E., Frischmuth, S., Holm, M., Becker, L., Flamm, K., Laszig, R., et al. (1999). Speech recognition, speech production and speech intelligibility in children with hearing aids versus implanted children. *International Journal of Pediatric Otorhinolaryngology*, 47, 165-169.

Ludlow, C. L. (2005). Central nervous system control of the laryngeal muscles in humans. *Respiratory Physiology & Neurobiology, 147*, 205-222.

Macken, M. A. & Barton, D. (1979). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, 7, 41-74.

Michalewski, H. J., Starr, A., Nguyen, T. T., Kong, Y-Y., & Zeng, F-G. (2005). Auditory temporal processes in normal-hearing individuals and in patients with auditory neuropathy. *Clinical Neurophysiology*, 116, 669-680.

Milenkovic, P. (2002). Time-frequency analysis software for 32-bit Windows (TF32) [Computer program]. Madison, WI: University of Wisconsin.

Miller, J. L., Green, K. P., & Reeves, A. (1986). Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast. *Phonetica*, 43, 106-115.

Nagao, K., & de Jong, K. (2007). Perceptual rate normalization in naturally produced rate-varied speech. *Journal of the Acoustical Society of America*, 121, 2882-2898.

Roman, S., Canevet, G., Lorenzi, C., Triglia, J-M., & Liegeois-Chauvel, C. (2004). Voice onset time encoding in patients with left and right cochlear implants. *Neuroreport*, 15 (4), 601-605.

Santarelli, R., Magnavita, V., De Filippi, R., Ventura, L., Genovese, E., & Arslan, E. (2009). Comparison of speech perception performance between Sprint/Esprit #G and Freedom processors in children implanted with Nucleus cochlear implants. *Otology and Neurotology*, 30, 304-312.

Sharma, A., Nash, A. A., & Dorman, M. (2009). Cortical development, plasticity and re-organization in children with cochlear implants. *Journal of Communication Disorders*, 42, 272-279.

Simonyan, K. & Jurgens, U. (2003). Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. *Brain Research*, 974, 43-59.

Stathopoulos, E. T., & Sapienza, C. M. (1997). Developmental changes in laryngeal and respiratory function with variations in sound pressure level. *Journal of Speech, Language, and Hearing Research*, 40, 595-614.

Theodore, R. M., Miller, J. L., & DeSteno, D. (2009). Individual talker differences in voice-onset time: Contextual influences. *Journal of the Acoustical Society of America*, 125 (6), 3974-3982.

Timgren, C. R., Campbell, M. & Boliek, C.A. (in preparation). Physiological and acoustic measures of speech and voice outcomes in children with cochlear implants. Unpublished master's thesis, University of Alberta, Edmonton, Alberta.

Tye-Murray, N., Spencer, L., & Gilbert-Bedia, E. (1995). Relationships between speech production and speech perception skills in young cochlear-implant users. *Journal of Acoustical Society of America*, 98 (5), 2454-2460.

Uchanski, R. M., & Geers, A. E. (2003). Acoustic characteristics of the speech of young cochlear implant users: A comparison with normal-hearing age-mates. *Ear & Hearing, 24*(1), 90S-105S.

Volaitis, L. E., & Miller, J. L. (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *Journal of Acoustical Society of America, 92*(2), 723-735.

Whiteside, S. P., Dobbin, R., & Henry, L. (2003). Patterns of variability in voice onset time: A developmental study of motor speech skills in humans. *Neuroscience Letters,* 347, 29-32.

Whiteside, S. P., & Marshall, J. (2001). Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica, 58*(3), 196-210.

Wohlert, A. B., & Hammen, V. L. (2000). Lip muscle activity related to speech rate and loudness. *Journal of Speech, Language, and Hearing Research*, 43, 1229-1239.

Woo, P. (1996). Quantification of videostrobolaryngoscopic findings – measurements of the normal glottal cycle. *Laryngoscope*, 106(3), 1-27.

Yalcinkaya, F., Muluk, N. B., Atas, A., & Keith, R. W. (2009). Random gap detection test and random gap detection test-expanded results in children with auditory neuropathy. *International Journal of Pediatric Otorhinolaryngology*, 73, 1558-1563.

Zhang, S. P., Bandler, R., & Davis, P. J. (1995). Brain stem integration of vocalization: Role of the nucleus retroambigualis. *Journal of Neurophysiology, 74*(6), 2500-2512.

Zlatin, M. A., & Koenigsknecht, R. A. (1976). Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research*, 19, 78-92.

Zwolan, T. A. (2009). Cochlear implants. In J. Katz, L. Medwetsky, R. Burkard, & L. Hood, *Handbook of Clinical Audiology (6th Edition)* (pp.912-933). Baltimore: Lippincott Williams & Wilkins.

Appendices

Appendix A: Recruitment Poster for Typically Developing Participants

The Effect of Voice Intensity and Speech Rate on Voice-Onset Time in Typically Developing Children and Children with Cochlear Implants

Investigators

Carol Boliek, PhD
Melanie Campbell, PhD
Erica Knuttila, BSc

We want to understand the relationship between how fast and loud children talk and how well they are understood by others. We are looking for children between the ages of 6 and 12 who are typically developing to serve as a comparison group to children who have cochlear implants. This study will help us understand the control of voice and speech in typical children in order to compare them to children with cochlear implants. This will hopefully lead to better voice and speech therapies that can be used with children who have cochlear implants.

If you would like to know more about our study, please contact

Dr. Carol Boliek, or Erica Knuttila at 780-492-0841 or Dr. Melanie Campbell at 780-492-0838

Thank you.

Appendix B: Recruitment Letter for Participants with Cochlear Implants

Project Title: The Effect of Voice Intensity and Speech Rate on Voice-Onset Time in Typically Developing Children and Children with Cochlear Implants

Investigators:
Supervisors: Dr. Carol Boliek and Dr. Melanie Campbell
Erica Knuttila, BSc; MSc-SLP candidate

Your child is being asked to take part in a research study because s/he has a cochlear implant, is between the ages of 5 and 12, was implanted before age three and is an oral communicator. As the parent, you will be asked to provide consent for your child's participation and allow us to obtain information about the implant from your child's health records at the Glenrose Rehabilitation Hospital. You may attend a one-time study with your child or have someone else bring your child in.

Study Information:

We want to know how speech develops in children with cochlear implants. We want to learn whether or not children with cochlear implants are able to adjust the fine movements of speech and voice in response to changes in speaking rate and loudness. During this study, Erica Knuttila will be recording and then measuring a timing feature of your child's speech. This timing feature is an important part of speech that allows us to tell the difference between speech sounds in words like pat/bat or tot/dot. As well, Erica will measure your child's speech understanding. We hope to better understand the control of voice and speech in children with cochlear implants in order to develop better voice and speech therapies for children with hearing loss. We hope to provide information that will help improve cochlear implant technology for children in the future.

You and your child will be invited to the University of Alberta at a time **most convenient** for your family (**including evenings and weekends**). Listening tasks will be done. Also audio recordings will be made of your child's voice for later analysis. The procedure will take approximately one hour to complete. There are no direct benefits to you or your child by taking part in the study. However, all of your child's results will be written up and summarized and provided to you upon completion of the study. To compensate your child for his/her time, a small gift and Marble Slab food coupon will be provided at the end of the study. We will pay for your parking at Corbett Hall. There are no risks to you or your child by taking part in the study. Nothing we do in this study will harm your child. Your child may chose to bring a friend to help make the experience more fun. The friend may also enroll in the study with parent consent. Participation in the study is voluntary. Your child will be free to discontinue the project at any time. You and your child do not need to give a reason. Services that you receive at the Glenrose Rehabilitation Hospital will not be affected.

Contact:

Thank you for taking the time to read this letter. If you are interested in participating in this one time study or have any further questions, please feel free to call Dr. Carol Boliek at 780-492-0841 or Dr. Melanie Campbell at 780-492-0838. You may email Erica Knuttila at knuttila@ualberta.ca, Dr. Carol Boliek at carol.boliek@ualberta.ca, or Dr. Melanie Campbell at melanie.campbell@ualberta.ca.

Appendix C: Recruitment Poster for Participants with Cochlear Implants

The Effect of Voice Intensity and Speech Rate on Voice-Onset Time in Typically Developing Children and Children with Cochlear Implants

Investigators

Carol Boliek, PhD
Melanie Campbell, PhD
Erica Knuttila, BSc

We want to understand the relationship between how fast and loud children talk and how well they are understood by others. We are looking for children between the ages of 6 and 12 who wear cochlear implants, were implanted before the age of three, are oral communicators and attend school in their local community. This study will help us understand the control of voice and speech in children with cochlear implants. This will hopefully lead to better voice and speech therapies that can be used with children who have hearing loss and help improve cochlear implant technology for children in the future.

If you would like to know more about our study, please contact

Erica Knuttila at knuttila@ualberta.ca, Dr. Carol Boliek at 780-492-0841 or Dr. Melanie Campbell at 780-492-0838

Thank you.

Appendix D: Parent Information Letter for Typically Developing Participants

Project Title: The Effect of Voice Intensity and Speech Rate on Voice-Onset Time in Typically Developing Children and Children with Cochlear Implants

Investigators:
Supervisors: Dr. Carol Boliek and Dr. Melanie Campbell
Erica Knuttila, BSc; MSc-SLP candidate

Your child is being asked to take part in a research study because he or she has normal hearing and is typically developing. As the parent, you will be asked to provide consent for your child's participation and attend the one-time study with your child.

Purpose of the study:
We want to know how speech develops in typical children. We want to learn whether or not typical children are able to adjust the coordination of speech and voice in response to changes in speech rate and vocal loudness. During this study, Erica Knuttila will be recording and then measuring a timing feature of your child's speech. This timing feature is an important part of speech that allows us to tell the difference between speech sounds in words like pat/bat or tot/dot. As well, Erica will measure your child's speech understanding. We hope to learn more about the control of voice and speech in typical children in order to compare it to that of children with hearing loss, leading to improved voice and speech therapies for them.

Procedure:
The researcher(s) will come to your home for testing or if you prefer, you can bring your child to the University of Alberta. A hearing screening and listening tasks will be done. Audio recordings will also be made. The procedure will take approximately one hour and fifteen minutes to complete.

We will check your child's hearing. He or she will put on headphones and be asked to listen for very soft beeping sounds, first in one ear and then the other. Your child will be asked to raise his or her hand when s/he hears a beep.

Then we will ask your child to listen to another set of beeps and clicks. We will ask your child to raise his/her hand if s/he hears two beeps/clicks.

Next, we will ask your child to listen to sentences. We will ask your child to point to and say the syllable s/he hears in each sentence.

We will ask your child to wear a microphone on his/her head. We will also ask your child to produce the phrase "It's a /*C*/od again" multiple times with any of p/b/t/d/k/g in the consonant "C" slot while speaking normally, twice as fast as normal, half as fast as normal, twice as loud as normal and half as loud as normal.

We expect the entire procedure to take approximately one hour and fifteen minutes. You and your child can take a break at any time. All of the tasks will be presented through game-like activities, which children typically enjoy. We will reimburse you for your parking fees, if necessary.

If you child becomes anxious or tired, we will stop and take a break. We will continue if and when your child says that he/she would like to carry on.

Confidentiality:
All of the information that we gather during this study will be kept confidential. Only the researchers directly involved in this study will know the identity of your child. The audio recordings of the session and any other related data will be locked in Dr. Melanie Campbell's lab at the University of Alberta. The tapes and data files will be labeled with a number code to keep all information private. We will not use these tapes for educational purposes unless we get your permission first. The data will be stored for at least five years per University of Alberta guidelines.

If requested, you will be informed of any publication of this study.

Benefits:
The results obtained from your child's hearing screening will be shared with you. You will be given a copy of these results for your records.

Risks:
There are no direct personal risks to taking part in this experiment. If the researcher(s) identify any problems from your child's hearing screening such as a hearing loss, we will refer your child to the appropriate professionals who are able to provide further testing and diagnosis.

Withdrawal:
Participation in the study is voluntary. Your child will be free to discontinue the project at any time. You and your child do not need to give a reason.

Contact:
Thank you for taking the time to read this letter. If you have any further questions, please feel free to call Erica Knuttila, Dr. Carol Boliek or Dr. Melanie Campbell at 780-492-0841. You may email Erica Knuttila at knuttila@ualberta.ca, Dr. Carol Boliek at carol.boliek@ualberta.ca, or Dr. Melanie Campbell at melanie.campbell@ualberta.ca.
Should you have any concerns about this study, you can contact Dr. Joanne Volden, Associate Dean, Graduate Studies and Research, Faculty of Rehabilitation Medicine at 780-492-0651.

Appendix E: Parent Information Letter for Participants with Cochlear Implants

Project Title: The Effect of Voice Intensity and Speech Rate on Voice-Onset Time in Typically Developing Children and Children with Cochlear Implants

Investigators:
Supervisors: Dr. Carol Boliek and Dr. Melanie Campbell
Erica Knuttila, BSc; MSc-SLP candidate

Your child is being asked to take part in a research study because s/he has a cochlear implant. As the parent, you will be asked to provide consent for your child's participation, allow us to obtain information about the implant from your child's health records at the Glenrose Rehabilitation Hospital and attend a one-time study with your child.

Purpose of the study:
We want to know how speech develops in children with cochlear implants. We want to learn whether or not children with cochlear implants are able to adjust the fine movements of speech and voice in response to changes in speaking rate and loudness. During this study, Erica Knuttila will be recording and then measuring a timing feature of your child's speech. This timing feature is an important part of speech that allows us to tell the difference between speech sounds in words like pat/bat or tot/dot. As well, Erica will measure your child's speech understanding. We hope to better understand the control of voice and speech in children with cochlear implants in order to develop better voice and speech therapies for children with hearing loss.

Procedure:
You and your child will be invited to the University of Alberta. Listening tasks will be done. Also audio recordings will be made. The procedure will take approximately one hour to complete.

We will ask your child to listen to beeps and clicks and to raise his/her hand if s/he hears two beeps/clicks.

Next, we will ask your child to listen to sentences and to point to and then say the syllable s/he hears in each sentence.

We will ask your child to wear a microphone on his/her forehead. We will also ask your child to produce the phrase "It's a /*C*/od again" multiple times with any of p/b/t/d/k/g in the consonant "C" slot. You child will be asked to speak normally, twice as fast as normal, half as fast as normal, twice as loud as normal and half as loud as normal.

We expect the entire procedure to take approximately one hour. You and your child can take a break at any time. All of the tasks will be presented through game-like activities, which children typically enjoy. We will reimburse you for your parking fees.

If your child becomes anxious or tired, we will stop and take a break. We will continue if and when your child says that s/he would like to carry on.

A speech-language pathologist at the Glenrose will provide the following information from your child's health record: his/her name, address, date of birth, date at implantation, type of cochlear implant, surgical insertion depth, length of cochlear implant use, and most recent hearing and speech test results. This information will be taken only from your child's current health record at the Glenrose Hospital. The information will be filed under your child's assigned participant number only and not by his or her name.

Confidentiality:
By signing the consent form, you give permission to the study staff to gain access to the information items listed above from your child's personal health information. That information will be kept confidential. Only the researchers directly involved in this study will know the identity of your child. The audio recordings of the session and any other related data will be locked in Dr. Melanie Campbell's lab. The tapes and data files will be labeled with a number code to keep all information private. We will not use these tapes for educational purposes unless we get your permission first. The data will be stored for at least five years per University of Alberta guidelines.

If requested, you will be informed of any publication of this study.

Benefits:
There are no direct benefits to you or your child by taking part in the study.

Risks:
There are no risks to you or your child by taking part in the study. Nothing we do in this study will harm your child.

Withdrawal:
Participation in the study is voluntary. Your child will be free to discontinue the project at any time. You and your child do not need to give a reason. Services that you receive at the Glenrose Rehabilitation Hospital will not be affected.

Contact:
Thank you for taking the time to read this letter. If you have any further questions, please feel free to call Erica Knuttila, Dr. Carol Boliek or Dr. Melanie Campbell at 780-492-0841. You may email Erica Knuttila at knuttila@ualberta.ca, Dr. Carol Boliek at carol.boliek@ualberta.ca, or Dr. Melanie Campbell at melanie.campbell@ualberta.ca. Should you have any concerns about this study,

you can contact Dr. Joanne Volden, Associate Dean, Graduate Studies and Research, Faculty of Rehabilitation Medicine at 780-492-0651.

Appendix F: Consent Form – Parent

Project Title: The Effect of Voice Intensity and Speech Rate on Voice-Onset Time in Typically Developing Children and Children with Cochlear Implants

Investigators:

Co-Supervisors: Dr. Carol Boliek & Dr. Melanie Campbell
Erica Knuttila, BSc; MSc-SLP thesis candidate
Contact Number: (780) 492-7588

	Yes	No
Do you understand that you are volunteering for your child to be in a research study?	☐	☐
Have you read and received a copy of the attached information sheet?	☐	☐
Do you understand the benefits and risks involved in your child's taking part in this research study?	☐	☐
Have you had an opportunity to ask questions and discuss this study?	☐	☐
Do you understand that your child is free to withdraw from the study at any time without penalty?	☐	☐
Has the issue of confidentiality been explained to you?	☐	☐
Do you understand who will have access to the information you provide?	☐	☐
Do you understand that the session will be audiotaped for the purpose of later analysis?	☐	☐
Do you consent to the audiotapes being used for educational purposes?	☐	☐

This study was explained to me by: ______________________________

I agree to have my child take part in this study.

____
Child's name (printed)

____
Parent Name (printed)

____
Signature of Parent

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to have his/her child participate.

______________________________ ______________________
Signature of Investigator Date

Appendix G: Assent Form – Typically Developing Participants

Project Title: The Effect of Voice Intensity and Speech Rate on Voice-Onset Time in Typically Developing Children and Children with Cochlear Implants

Investigators:

Co-Supervisors: Dr. Carol Boliek and Dr. Melanie Campbell
Erica Knuttila, BSc; MSc-SLP thesis candidate

Why have you been asked to do this?
You have normal hearing. We want to find out how speech develops in children with or without normal hearing. We want to compare your speech to the speech of children who do not have normal hearing.

How long will this take?

It will take approximately 30 minutes to check your hearing and listening and 45 minutes to do the recording. You will be finished in one hour and fifteen minutes.

What will you have to do?
At the beginning, we will check your hearing. You will put on headphones. You will listen for very soft beeping sounds, first in one ear and then the other. You will raise your hand when you hear a beep. Checking your hearing will take about 10 minutes.

You will put on headphones for two more listening activities. First you will listen for another set of beeps. You will raise your hand each time you hear two beeps. Next you will listen to sentences. You will be given a piece of paper with six words on it. You will point to the word that you hear in each sentence. You will also tell us what you hear in each sentence.

You will then put on a microphone, so that we can record your voice onto a tape recorder.
You will say silly sentences like: "It's a Pod again", "It's a Bod again", "It's a Tod again", "It's a Dod again", "It's a Cod again", and "It's a God again" in your normal voice, five times.

Second, you will say the same phrases twice as fast as normal, five times.

Third, you will say the phrases half as fast as normal, five times.

Then, you will say the phrases twice as loud as normal, five times.

Finally, you will say the phrases half as loud as normal, five times.
The recording will take about 45 minutes. You can have a break anytime.

Will it help?

This project will help us find out how children with or without normal hearing learn to talk.

Will it hurt?

Nothing we are asking you to do will hurt. Nothing will be hard for you to do.

Can you quit?

You don't have to take part in the study at all. You can quit at any time. If you want to quit, you can tell your parents or the researchers.

Who will know?

Only your parents and the researchers will know you're taking part in the study unless you want to tell someone. Only the researchers will know your name and your information.

Your signature:

If you want to take part in this study, please sign your name below to show that you agree to take part.

Do you have any questions?

You can ask the researchers or your mom or dad about any part of this study at any time.

I agree to take part in the study.

_____________________________ _____________________

Signature of Research Participant Date

____________________________ _____________________

Signature of Investigator Date

Appendix H: Assent Form – Participants with Cochlear Implants

Project Title: The Effect of Voice Intensity and Speech Rate on Voice-Onset Time in Typically Developing Children and Children with Cochlear Implants

Investigators:

Co-Supervisors: Dr. Carol Boliek and Dr. Melanie Campbell
Erica Knuttila, BSc; MSc-SLP thesis candidate

Why have you been asked to do this?

You have a cochlear implant. We want to find out how children with cochlear implants learn to talk.

How long will this take?

It will take approximately one hour to finish recording.

What will you have to do?

We will check your listening. You will listen for beeps coming from the computer speakers. You will raise your hand each time you hear two beeps.

You will listen to sentences through computer speakers. You will be given a piece of paper with six words on it. You will point to the word you hear in each sentence. You will also tell us what you hear in each sentence.

Next, you will wear a tiny microphone on your forehead the whole time, so that we can record your voice onto a tape recorder.

First, you will say silly sentences like: “It’s a Pod again”, “It’s a Bod again”, “It’s a Tod again”, “It’s a Dod again”, “It’s a Cod again”, and “It’s a God again” in your normal voice, five times.

Second, you will say the same phrases twice as fast as normal, five times.

Third, you will say the phrases half as fast as normal, five times.

Then, you will say the phrases twice as loud as normal, five times.

Finally, you will say the phrases half as loud as normal, five times.
You can have a break anytime.

Will it help?

This project will help us find out how children who have cochlear implants learn to talk.

Will it hurt?

Nothing we are asking you to do will hurt. Nothing will be hard for you to do.

Can you quit?

You don't have to take part in the study at all. You can quit at any time. If you want to quit, you can tell your parents or the researchers.

Who will know?

Only your parents and the researchers will know you're taking part in the study unless you want to tell someone. Only the researchers will know your name and your information.

Your signature:

If you want to take part in this study, please sign your name below to show that you agree to take part.

Do you have any questions?

You can ask the researchers or your mom or dad about any part of the study at any time.

I agree to take part in the study.

______________________________ ____________________
Signature of Research Participant Date

______________________________ ____________________
Signature of Investigator Date

Appendix I: Picture description of VOT using Praat acoustic analysis software (Boersma & Weenink, 2010)

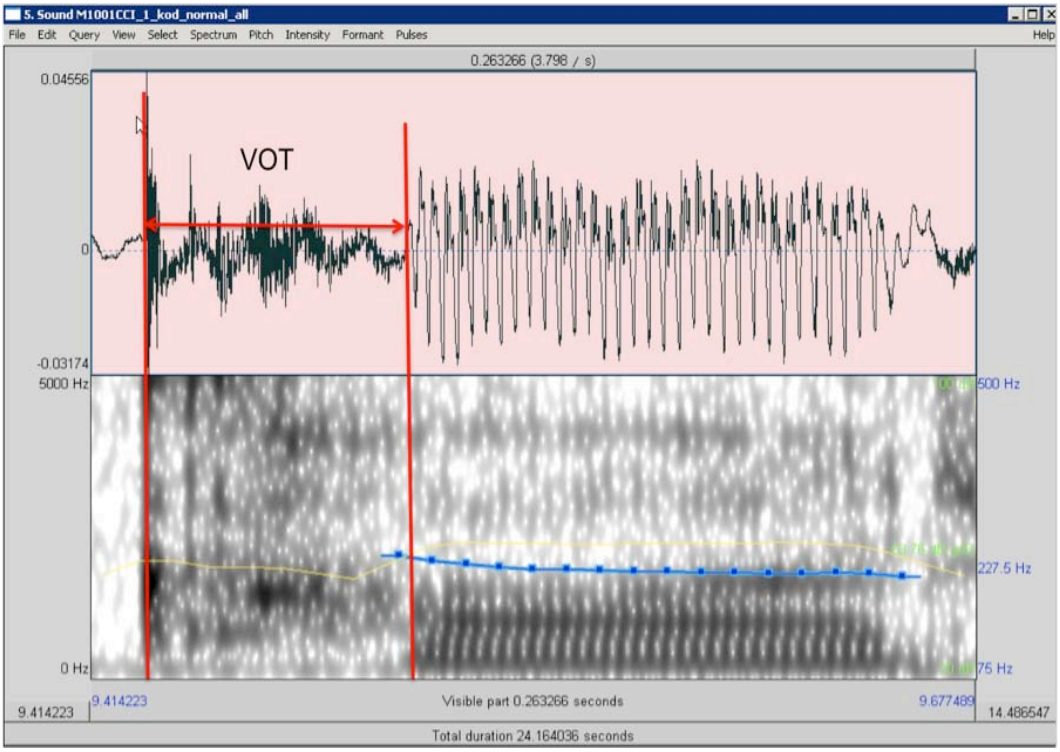

Figure I.1*: Picture description of VOT using* Praat *acoustic analysis software (Boersma & Weenink, 2010). Cursor on the left side indicated the burst of the stop consonant and the cursor on the right side indicated the onset of voicing for the vowel. The time between the two cursors was calculated as the VOT.*

Appendix J: Group medians for all conditions

Table J.1*: Group medians for all conditions*

Condition	**Median**
Voiced tokens, slow rate	18.3356667
Voiced tokens, soft voice	29.3057748
Voiced tokens, fast rate	16.0729063
Voiced tokens, loud voice	10.6309685
Voiced tokens, normal rate and loudness	16.9429448
Voiceless tokens, slow rate	106.7996667
Voiceless tokens, soft voice	85.9856948
Voiceless tokens, fast rate	40.1865929
Voiceless tokens, loud voice	42.898
Voiceless tokens, normal rate and loudness	73.0652556

Appendix K: Box and whisker plots, skewness, kurtosis and standard error values for the normal hearing group

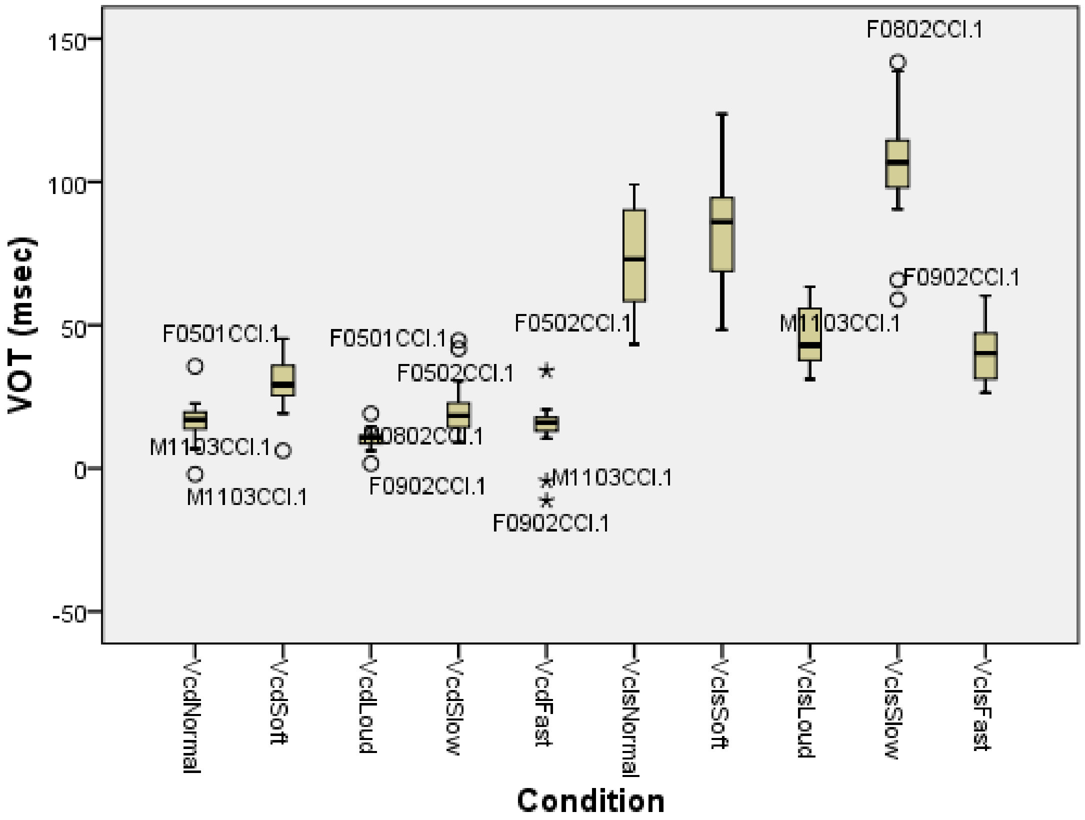

Figure K.1: *Group medians across conditions*

Table K.1: *Skewness and kurtosis across all conditions*

	Skewness		Kurtosis	
	Statistic	Std. Error	Statistic	Std. Error
VcdNormal	.009	.580	2.837	1.121
VcdSoft	-.525	.580	1.052	1.121
VcdLoud	.031	.580	1.686	1.121
VcdSlow	1.284	.580	1.117	1.121
VcdFast	-1.006	.580	2.627	1.121
VclsNormal	-.134	.580	-1.388	1.121
VclsSoft	.215	.580	-.378	1.121
VclsLoud	.335	.580	-1.224	1.121
VclsSlow	-.497	.580	.588	1.121
VclsFast	.291	.580	-1.139	1.121

Table K.2a: *Mean, standard error and 95% confidence intervals for voicing.*

1. Voicing

Measure:MEASURE_1

Voicing	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	18.339	1.829	14.416	22.262
2	69.294	3.322	62.170	76.419

Table K.2b: *Mean, standard error and 95% confidence intervals for condition.*

2. Condition

Measure:MEASURE_1

Condition	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	44.485	3.205	37.611	51.358
2	55.909	3.401	48.615	63.202
3	28.067	1.666	24.493	31.640
4	63.198	3.417	55.868	70.528
5	27.426	2.052	23.024	31.827

Table K.2c: *Mean, standard error and 95% confidence intervals for voicing x condition interaction.*

3. Voicing * Condition

Measure:MEASURE_1

Voicing	Condition	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	16.438	2.101	11.932	20.944
	2	29.704	2.575	24.181	35.226
	3	10.225	1.029	8.017	12.433
	4	21.316	2.687	15.552	27.080
	5	14.014	2.697	8.229	19.799
2	1	72.532	4.913	61.994	83.070
	2	82.114	5.605	70.092	94.136
	3	45.909	2.709	40.099	51.718
	4	105.080	5.810	92.618	117.542
	5	40.837	2.776	34.884	46.791