January 2014

Accuracy of /t/ Productions in Children with Cochlear Implants as Compared to Normal-Hearing, Articulation Age-Matched Peers

Terry Gier
University of South Florida, terry2@mail.usf.edu

Accuracy of /t/ Productions in Children with Cochlear Implants as Compared to Normal-Hearing, Articulation Age-Matched Peers

by

Terry Gier

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science
Department of Communication Sciences and Disorders
College of Behavioral and Community Sciences
University of South Florida

Major Professor: Ruth Huntley Bahr, Ph.D.
Theresa Chisholm, Ph.D.
Stefan Frisch, Ph.D.

Date of Approval:
July 21, 2014

Keywords: Speech production, Speech sound development, Cochlear Implant, Covert Contrast, Listener Ratings.

Acknowledgments

The successfully completion of this project would not have been possible without the efforts of several people:

I will be eternally grateful to Dr. Ruth Bahr for the feedback, support, and encouragement she gave. Our long conversations were instrumental in developing the idea for this project, and her words of encouragement were essential in its completion.

I would also like to thank Dr. Stefan Frisch for his insights and flexibility and Dr. Theresa Chisolm for her support throughout this process.

I also owe thanks to Dr. Jean Krause for her last-minute participation that allowed this project to move forward.

Finally, I must thank my family who gave unconditional support and love when I needed it most. I would never have made it to this point without you.

Table of Contents

List of Tables

List of Figures

Abstract

Investigations focused on speech sound development in children with CIs have shown initial accelerated growth, followed by a plateau where consonant order of acquisition generally mirrors that of NH children, but is slower (Blamey, Barry, & Pascale, 2001; Serry & Blamey, 1999; Spencer & Guo, 2013). A notable exception to this pattern, /t/, has been shown to be acquired later-than normal in several investigations (Blamey et al., 2001; Chin, 2003; Ertmer, True Kloiber, Jongmin, Connell Kirleis, & Bradford, 2012). The primary purpose of this investigation was to 1) examine the accuracy of /t/ productions in children with CIs and 2) quantify subtle phonetic differences in correctly produced consonants and substituted consonants (or covert contrast).

Two groups of children, who had participated in a larger study that examined the influence of speech production abilities on speech perception scores in children with CI (Gonzalez, 2013), provided the speech stimuli for this investigation. The experimental group included nine congenitally deafened children with CI, ranging in age from 2;11 to 6;4 years (M=4;9), who were implanted by 3 years of age, had at least 12 months of device experience, and only used an oral mode of communication. These children were matched to typically developing children by articulation ability and gender.

Recordings of the verbal responses on the OlimSpac were obtained from the Gonzalez (2013) study. Thirty-three graduate students in speech-language pathology used a 7 point equal-appearing interval scale to rate the phonetic accuracy of /t/ and the phonemes that were found to be most often substituted for it, /d/ and /□/.

A three-way ANOVA was performed to determine the differences in perceived consonant accuracy across: group, transcription category, and phoneme substitution. The significant interaction between group and transcription category was of particular interest. Results indicated that children with CIs did not show an unusually delayed development of /t/. When a confusion matrix was generated to depict overall OlimSpac performance, the NH group was noted to outperform the CI group across all phonemes. This would suggest that /t/ was not uniquely poorer in the CI group, but instead these children evidenced poorer phoneme accuracy in general. Finally, group differences also were apparent in substitutions of [t] for target /d/ and /□/ productions (i.e., covert contrast). Clinical implications are described.

Chapter 1

Literature Review

Cochlear Implants (CIs) have been approved for use in children over 2 years of age since 1989 and in children as young as 12 months since 2000 (National Institute of Health and Human Services, 2010). Early implantation has been strongly advocated and repeatedly shown to be beneficial since then as a means to encourage timely and accurate development of speech (Shu-Chen, Spencer, & Tomblin, 2004; Tomblin, Barker, Spencer, Xuyang, & Bruce, 2005; Tye-Murray, Spencer, & Woodworth, 1995). However, even when implanted relatively early, children with CI still have lost the early hearing experiences available to all children with normal hearing (NH), which may significantly impact their speech development. For example, Stoel-Gammon and Otomo (1986) reported that although NH infants and infants with hearing loss both began babbling at the same age, their speech development did not follow the same trajectory. These authors found that over a 12 month period, the variety of consonants produced by children with hearing loss either plateaued or decreased. Conversely, the consonant inventories of children with NH increased significantly and these differences were noted, even in the babbling of children at 8 months of age. Hence, despite timely restoration of auditory input, deaf children with cochlear implants are still decidedly at a disadvantage.

Maturity of the speech mechanism and motor cortex, knowledge of the acoustic, perceptual and articulatory characteristics of speech sounds, knowledge of the phonotactic rules of a language (Munson, Edwards & Beckman, 2005) and, perhaps most importantly, hearing acuity are all necessary prerequisites to the development of speech in children. Yet children are proven masters when tasked with deciphering the arbitrary rules and rhythms of their native language. Despite the complexity inherent to speech, children progress from random babbling behaviors to the production of words quickly, usually within 6-8 months (American Speech-Language-Hearing Association, n.d.). Given this natural aptitude, it is no surprise that deaf children go on to develop speech nearly as well as children with NH when implanted early. Although children with CIs develop remarkably accurate speech in a timely manner, several differences from NH children exist. What follows is a review of the current literature regarding speech sound acquisition in children with CIs. First, patterns of speech sound development in children with CIs will be described. Next, neurocognitive differences between CI recipients and the psychoacoustic effects of the altered CI signal will be explored as explanations for the noted variability in speech outcomes for children with CIs. Finally, a more sensitive approach to the perceptual assessment of speech sound accuracy will be described.

Patterns of Speech-Sound Acquisition in Children with CIs

Once implanted, children with hearing loss have been shown to resume a normal, or near-normal developmental rate and sequence of speech sound acquisition (Spencer and Guo, 2013). In fact, several researchers have demonstrated that no significant differences exist between the speech production abilities of children with CIs and typically developing children when device experience, as opposed to chronological age, was used as a referent (Flipsen, 2011; Spencer & Guo, 2013). Specifically, Spencer and Guo (2013) reported that 50% of early-implanted children with one year of device experience scored within the normal range for chronological age on the *Goldman-Fristoe Test of Articulation, Second Edition* (GFTA-2; Goldman & Fristoe, 2000).

Because the children with CIs were able to acquire language in 12 months that children with NH developed over 2 years, this finding strongly suggests that children with CIs are able to "catch up" much more quickly than NH children with similar language experience. Nevertheless, differences in the order and rate of consonant acquisition between CI and NH children have been reported.

Findings regarding consonant inventory and rate of acquisition. Several authors have attempted to measure consonant inventory size in children with CIs at different stages of development. For instance, Serry and Blamey (1999) followed nine Australian children with CIs (mean age of implantation = 3years, 9 months) for four years. These children were videotaped during natural play with a familiar adult six and three months prior to implantation and bi-annually thereafter. The conversations produced by these children were subsequently analyzed using the following criterion. A phoneme was considered acquired by a child if at least 50% of their attempts were produced correctly in intelligible words. The phoneme was considered mastered by the group if at least five of the nine children had acquired it. After four years of robust hearing experience, the children with CIs only mastered 13 of 24 phonemes. Glides and nasals were acquired earliest. Fricatives and affricates were acquired last. Additionally, the more "visible" bilabial and labio-dental sounds were acquired earlier than other consonants. These findings are generally consistent with the order of acquisition for typically developing children. The rate, however, was slower since typically developing children had acquired 22 of 24 consonants at four years of age. In a follow-up to this study, Blamey, Barry and Pascale (2001) discovered that of the 11 consonants that were not present in the children's inventories after four years, six were still not evident after six years: /t, s, □, z, θ, □/. They interpreted this result as suggesting an initial burst of consonant inventory growth, followed by a period of slower development.

These acquisition findings were confirmed by Chin (2003), who analyzed the speech of 12 children who had five years of CI experience. Speech was elicited in single words using a picture naming task. The test probed all English consonants “in three word positions, in several different words, and under conditions of potential morphophonemic alternations” (p. 852). Consistent with the results of the previous studies (Blamey et al., 2001; Serry and Blamey, 1999), these children showed the greatest difficulty with producing alveolar and interdental fricatives. Unlike the children in the previous studies, however, these children (as a group) had mastered affricates by five years post-implantation.

More recent studies suggest that order of consonant acquisition in children with CIs differs slightly from that of typically developing children. Ertmer, et al. (2012) used a sentence repetition task to elicit speech from 11 young children with 24 months of CI experience (mean age of implantation = 17.6 months) and their age- and gender-matched peers. Their results for initial consonants generally agreed with the Iowa-Nebraska norms (Smit, Hand, Freilinger, Bernthal, & Bird, 1990) for both groups of children, with phonemes that are recognized as early-developing produced more accurately than middle- or late-developing phonemes (Shriberg, Gruber & Kwiatkowski, 1994). The results for final consonants, however, did not follow the same pattern. Children with NH produced early sounds most accurately, followed by middle and late sounds. Children with CIs were much less accurate producing all final consonants, regardless of developmental sequence. Interestingly, the children with CIs were also more accurate when producing the late sounds (when in the final position) than either the early or middle sounds. The authors noted that this may indicate that the children with CIs had actually acquired the “late” sounds first.

Spencer and Guo (2013) also utilized the Iowa-Nebraska norms (Smit et al., 1990) to compare the speech of children with CIs and typically developing children. In this study, 32 children who were implanted prior to 30 months old were tested with the GFTA-2 on the anniversary of the activation of their cochlear implant (±3 weeks) for four consecutive years. Results revealed that by 24 months of device experience, the mean GFTA-2 standard score achieved by children with CIs was within normal limits. The order of consonant acquisition was also similar to that of typically developing children (anterior before posterior sounds, stops before fricatives, etc.). The authors noted, however, that several sounds were acquired earlier by the CI children (when matched for hearing experience) than the children in the Smit et al. (1990) normative sample. These sounds included: /f, s, □/ in the initial position and /l, s, f, □/ in the final position, lending credence to earlier findings of atypical order of acquisition for select phonemes. When rate of consonant acquisition was considered, these researchers reported an early burst of improvement (between 24-36 months) followed by a period of slower development (between 36-48 months), which is consistent with previous studies (Blamey et al., 2001, Chin, 2003, Ertmer et al., 2012).

Findings regarding speech accuracy and error types. Many of these studies also investigated the nature of the most common errors in CI children’s speech. Blamey et al. (2001), using narrow transcription, determined that the majority of errors involved in sounds not yet considered acquired involved distortions (i.e., unaspirated /t/, or lateralized /s/). The next most common error type was substitutions, followed by deletions. Careful inspection of these results revealed that several delayed phonemes shared the alveolar/palatal-alveolar place of articulation.

Because several late-acquired sounds (such as /t, and □/) shared a common place of articulation, the authors hypothesized that lack of fine motor control for place of articulation may be causing those sounds to be unusually delayed. However, this idea does not seem to hold when looking at the data as a whole, since place errors did not predominate overall.

When using broad instead of narrow transcription, Spencer and Guo (2013) found error patterns more consistent with those seen in typically developing children. They noted that substitutions were more common in the initial position of words and omission errors were more common in the final position. Omission errors remained high for the first three years of implant experience and were noted to drop in the fourth year.

Where previous studies described the acquisition of English consonants, Chin (2003) noted that it was important to describe all speaker productions, because successful acquisition of a language also involves exclusion of non-native sounds. Where NH children do not typically produce non-native sounds, children with CIs do so frequently. Using narrow transcription, Chin (2003) analyzed participants' speech for the presence of non-English sounds in addition to the more common search for missing segments. He found that several consonants only differed slightly and were affected by distortion-like processes, such as devoicing or dentalization. Some consonants, however, were substituted with sounds completely alien to the English language, including a velar fricative, a voiceless uvular stop and a dental stop. Teoh and Chin (2009) assert that these non-typical errors are common amongst children with CIs, and that these errors often make their speech sound accented. These authors argue that clinicians must not focus only on general intelligibility, but must also ensure that children with CIs produce all native sounds (including appropriate allophonic variations) correctly. They stress that it is sometimes more difficult for children with CIs to seamlessly identify with their ambient social group if their speech is anomalous, or sounds like a “foreign accent”.

Atypical emergence of select consonants. Rather than following a strictly developmental sequence, the speech of children with CIs has been shown to deviate in several important ways. For example, several typically later-developing phonemes were found to emerge earlier. These included /r/ (Ertmer et al., 2012), and /□/ (Ertmer et al., 2012; Spencer et al. 2013), among others. However, /□/ did not emerge early in some other studies (Blamey et al., 2001; Chin, 2003; Serry et al.,1999) and /t/, which typically develops early in NH children, was repeatedly found to be delayed in children with CIs (Blamey, et al., 2001; Chin, 2003; Ertmer et al., 2012; Spencer & Guo., 2013). According to the Iowa-Nebraska norms (Smit et al., 1990), 87% of typically developing children produce /t/ (or a phonotactically appropriate allophone) at 36 months. However, in children with CIs, the acquisition of /t/ was noted 4 to 6+ years post-implantation. This low level of accuracy could not be related to place of articulation since production of both /d/ and /n/ had very high accuracy rates (each above 80%) in the initial position of words. So, the difficulties with /t/ production are not easily explained.

Overall, this body of research clearly showed that CI participants made steady progress toward a complete English consonant inventory. Although Spencer and Guo (2013) asserted that the proportion of distortion errors was very low (<1%) when errors were classified using broad transcription, other studies that used narrow transcription (Blamey et al., 2001; Chin, 2003) found a large proportion of errors involving allophonic variations of the target sound. This finding would indicate that children with CIs are making many errors that are very close to the target production, or that they are gradually acquiring adult-like speech. However, characteristic of what is widely known about the speech development of children with CIs, there was significant variability seen both between and within previously discussed findings.

The Effects of Neurocognitive Differences on CI Outcomes

Children with CIs acquire speech similarly to their NH peers with some exceptions, but there remains large (and seemingly unpredictable) variability in individual outcomes. This is one of the most oft-cited frustrations of researchers and clinicians who work with children with CIs. While variability is most often attributed to age of implantation and length of pre-implantation deafness (Wilson & Dorman, 2008), other factors have been found to explain outcome variability in children with CIs: communication mode, clinic location (which impacts selection, management and assessment of research subjects), and speech processor. However, at least a third of variability between subjects is unaccounted (Sarant, Blamey, Dowell, Clark, & Gibson, 2001). Perhaps this some of this unexplained variability is neurocognitive in origin.

A cognitive explanation of variability is most apparent when examining the outcomes of children who are ostensibly implanted in identical (or nearly so) circumstances. Prelingually deaf children who have identical etiologies, who are implanted at the same age, and who receive identical therapies, will often not develop speech uniformly. The “Pisoni stars” are perhaps the best illustration of this common phenomenon (Pisoni, Cleary, Geers & Tobey, 1999). These researchers selected 27 prelingually deafened children from a larger pool of children with CIs who had previously been tested as part of a different investigation. These children were given the *Phonetically Balanced Kindergarten Word* (PBK) test, which is a difficult open-set word recognition task. The “Pisoni Stars” achieved scores in the top 20% of all children with CIs who had taken the PBK test and were then were compared to 23 children who achieved scores in the lowest 20% and who had the same amount of implant experience (i.e., two years). Since the children did not differ significantly on preimplant measures of cognition (such as non-verbal intelligence and attention), cognitive differences were judged to have no predictive value on post-implantation PBK test scores. However, results did reveal correlations between the PBK results and several other related measures, such as speech perception, language comprehension, spoken word recognition, and speech intelligibility, only in the high-performing group. These findings strongly suggest that there is a common underlying source of variance, generally attributable to biological, neurologic differences among children that affect learning and memory. Because natural cognitive linguistic aptitude is not exclusive to children with normal hearing, some deaf children are undoubtedly inherently more skilled at the language acquisition task.

Further, individual differences in brain plasticity have been shown to greatly impact speech perception and acquisition outcomes. This is due to a process called “cross-modal plasticity” that occurs following extended periods of sensory deprivation when other areas of the cortex (usually neighboring sensory modalities) begin to encroach upon the suddenly inactive cortical area (Wilson & Dorman, 2008). In other words, when the auditory cortex is “silent” due to prolonged deafness, cortical areas devoted to processing visual or somatosensory information begin to “take it over”. Naturally, the more cross-modal plasticity that occurs following the loss of a sensory modality (such as hearing), the harder it will be for the brain to interpret information once sensory perception is restored. For instance, very young children have been shown to experience more favorable outcomes following implantation than older children. Wilson and Dorman (2008) contend that this demonstrates that either they are able to reorganize the sensory processing that takes place in the temporal lobe more quickly and more effectively, or that their brains have experienced less cross-modal plasticity before the activation of the CI.

Seeking to support the idea of a neural-based explanation for variability in outcomes, Sharma, Dorman and Spahr (2002) compared the cortical responses to speech sounds of early (before age 4) and late-implanted (after age 4) children with identical device experience. Cortical responses (or latencies) are very short in NH children, which is indicative of normally-functioning temporal lobes. Although both groups of children with CIs initially demonstrated longer-than-normal latencies as expected, the latencies of the early-implanted group improved faster and to a larger degree than the later-implanted group. In fact, the early-implanted group attained latencies in the normal range after 5 months of CI use, on average. The later-implanted group did not achieve normal latencies, even after one year of use. These results demonstrated the younger brain's superior ability to respond and adapt to altered sensory input.

Further support of this idea was provided by analysis of pre- and post-implantation positron emission tomography (PET) brain scans of ten prelingually deafened children (Lee et al., 2001). The children (ranging in age from 2-20 years) had similar post-implantation device experience when their speech perception was later measured using a sentence recognition task. These authors found that amount of device experience did not explain the variability seen between subjects. Rather, the results revealed that the source of variance was differences in the children's neural characteristics. The children who experienced the most success post-implantation had lower than normal activity in the auditory cortex when it was measured prior to implantation. Those who showed this pre-implantation cortical hypoactivity tended to be younger and had shorter durations of deafness, supporting the idea that more cross-modal plasticity occurs after longer durations of deafness. Hence, longer periods of sensory deprivation will cause other sensory modalities to encroach upon the areas of the cortex normally assigned to receive the missing input and that this process is not readily reversible once input is restored (Lee, et al., 2001). Taken together, these findings reinforce a neurocognitive explanation for variance among cochlear implant users, be it physiologic or cognitive-linguistic.

Limitations of the CI Signal

Although the brains of certain CI users are better able to adapt to and make use of an impoverished signal, their perception abilities still usually fall short of "typical" (Caldwell & Nittrouer, 2013; Giezen, Escudero, & Baker, 2010). This is in part because the signal delivered by CIs, although adequate for reasonably accurate speech perception, is significantly degraded in relation to what is available to a person with NH (Pisoni et al., 1999; Spencer, 2002) due to the methods of speech processing common to all CIs. To be specific, speech processors preserve the amplitude envelope of speech, but much of the spectro-temporal "fine structure" information in speech is not transmitted (Wilson & Dorman, 2008). Since listeners must attend to both subtle temporal and spectral cues when distinguishing consonants, CI listeners are decidedly at a disadvantage.

The implications of this reduction in the transmission of speech sound information in children with CIs are especially pressing when one considers that children tend to weight acoustic properties of the speech signal differently than adults. A large body of research has shown that when attempting to interpret speech, children focus their attention on spectral cues, like formant transitions, to a higher degree than adults (Hicks & Ohde, 2005; Nittrouer & Burton, 2001). Perhaps relatedly, children have also been shown to develop a mature perception of another spectral characteristic of speech called "global spectral structure" by seven years of age, but do not display similar adult-like perception of amplitude envelopes (as produced by most CIs) by the same age (Nittrouer, Lowenstein, & Packer, 2009). Global spectral structure refers to the gross frequency characteristics of speech, and frequency changes over relatively long periods of time. The authors attribute the development of the ability to use global spectral structure to language experience rather than maturity of the central nervous system or speech mechanism since Mandarin speaking adult subjects, who have unquestionably mature auditory and motor pathways, also performed poorly on the same tasks.

Thus, because children with CIs resemble younger children in language experience, and older children in anatomical and neurologic maturity, it is possible that children with CIs attend to spectral information (including formant transitions) more than amplitude information in the same way as younger children with less language experience. However, like the amplitude envelope speech used by Nittrouer et al. (2009), the output of CI speech processors is very much lacking in all but the most crude spectral information (Wilson & Dorman, 2008). Rather than shift their attention to other acoustic cues that are salient to more experienced language users (amplitude information), children with CIs maintain a spectral cue weighting preference, even though they are not able to utilize spectral cue weighting as effectively as NH children (Giezen et al., 2010).

Research has shown that children with CIs maintain a spectral cue weighting strategy preference despite impoverished access to spectral information, but it has not directly established what effect this has on speech perception. However, indirect conclusions can be drawn based on the Giezen et al. (2010) data. For example, the children in this study performed most poorly on the place distinction task (a fu/su contrast). The salience of formant transitions (a spectral cue) in the perception of place of articulation of consonants has been well-established in the speech acoustics literature (Hicks & Ohde, 2005). It is also well-established that place of articulation is the most difficult phonetic feature to transmit via CI in both children and adults (Clark, 2003; Giezen et al., 2010; Pisoni et al., 1999). This pattern of deficit is consistent with the idea that children with CIs are not successful in discerning spectral acoustic cues, even when they are preferentially weighted. Conversely, manner and voicing are both consistently better transmitted by CIs, and perceived by CI users (Clark, 2003; Giezen et al., 2010; Pisoni et al., 1999). This is not surprising, given that the most salient acoustic cues related to these features are temporal in nature (voice onset time, continuance, stop-burst, etc.) and not spectral.

However, findings regarding perception of consonant features by children with CIs have been inconsistent. A recent study comparing early-implanted children and NH children matched for articulation ability using the GFTA-2 revealed counterintuitive results (Gonzalez, 2013). All children participated in the *On-Line Imitative Test of Speech-Pattern Contrast Perception* (OlimSpac; Boothroyd, Eisenberg, & Martinez, 2006) which required them to imitate several VCV non-words. A trained listener identified the consonant produced using a forced-choice task in order to ascertain the children's ability to perceive several phonetic contrasts, including place, manner and voicing. Gonzalez (2013) found that the NH participants most often failed the post-alveolar consonant place contrast and children with CIs most often failed the consonant voicing contrast. While these results do not support the idea of difficulty perceiving spectral cures, it may be that the children with NH were younger and still developing the post-alveolar place contrast, while the children with CI had possibly worked on this contrast in their speech therapy.

On the other hand, certain young CI users may uniquely excel at discerning spectral cues, and therefore be better able to perceive and therefore acquire speech. This could explain the previously "mysterious" success of a small percentage of children with CIs (Pisoni et al., 1999).

Research has demonstrated that a CI user's ability to weight acoustic cues similarly to NH adults is a prerequisite to highly accurate word recognition (Moberly et al., 2014). Specifically, these researchers showed that CI users' ability to use spectral cues predicted word recognition scores better than age of implantation, which was previously thought of as the variable that affected outcomes the most. Although spectral cue weighting has not been examined in depth in children with CIs, it is likely that successful CI use in children is associated with similar cue weighting strategies as adults.

While children do approach the language learning task individually, researchers also need to consider their own measurement techniques. Previous researchers differed in whether they used broad (Blamey et al., 2001; Ertmer et al., 2012; Spencer & Guo., 2013) or narrow (Blamey et al., 2001; Chin, 2003; Serry& Blamey, 1999) transcriptions and this seemed to affect the types of errors they noted. Given that children with CIs may use different acoustic cues when analyzing new phonemes, it is also possible that they display this knowledge in their speech sound productions. One way to begin to assess their productive knowledge is to have listeners rate the adequacy of their speech sound productions. This approach will be described next.

Listeners' Perception of Accuracy, and an Approach to Measurement

All prior descriptions and analyses of speech development in children with CIs used phonetic transcription. Specifically, some investigators utilized broad phonetic transcription (Ertmer et al., 2012; Spencer & Guo, 2013) which attempts to classify speech sounds into finite categories, and others used narrow phonetic transcription, which allows for the subtle variations that exist in all examples of natural speech production (Chin, 2003; Serry& Blamey, 1999). Narrow transcription, although able to convey more information regarding speech than broad transcription, is still inadequate when attempting to record gradual changes in consonant quality as children develop speech. For example, when children are beginning to develop speech, they often collapse phonemic contrasts into what is perceived as a single phoneme. However, upon acoustic analysis, there is often a statistically significant difference between two seemingly identical productions (Weismer, 1984). Children who make these distinctions are generally said to display "*covert contrast*" (Macken & Barton, 1980; Munson et al,, 2005). The contrast is "covert", or hidden, because it is not immediately obvious to listeners, and the two contrastive sounds are typically transcribed with the same character from the International Phonetic Alphabet. Acoustic analysis of children's speech has revealed that covert contrast is ubiquitously present in children's speech as they gradually acquire adult-like contrasts between speech sounds (Munson, Edwards, Schellinger, Beckman, & Meyer, 2010; Schellinger, Edwards, & Munson, 2010; Weismer, 1984). Covert contrast also has been demonstrated in the sound substitutions of children with many language backgrounds, who have typically developing speech and who have sound disorders (Li, Edwards, & Beckman, 2009; Scobbie, Gibbon, Hardcastle, & Fletcher, 2000).

While covert contrast is not present in all children, Munson, Schellinger, & Carlson (2012) assert that assessing the presence of this process is clinically important for three reasons. First, its absence has been strongly linked to increased treatment times in children with speech sound disorders (Tyler, Figurski, & Langdale, 1993). If clinicians can accurately determine the presence or absence of covert contrast, they can make more accurate prognoses, and make more informed decisions regarding therapy targets. Second, if a child displays covert contrtast, a clinician could rule out therapy choices that assume the child has no knowledge of the phoneme that is substituted (for example, that "bear" and "pear" should sound different). Third, the ability to measure the presence of covert contrast would imply the ability to track the progress of phoneme development from immaturity to maturity. Clinician would not be forced into a choice of either correct or incorrect, but would be able to track subtle changes as the child develops.

Given its clinical significance, it is important that clinicians are able to assess covert contrast quickly and reliably. However, all previous studies that investigated covert contrast did so using acoustic analysis, which is not always a clinically practical tool. Recognizing a clear need for research, Schellinger et al. (2010) sought to determine whether or not listeners could distinguish fine phonetic detail in children's speech using another means of measurement, visual analogue scales (VAS). They presented listeners with children's productions of /s/ and /θ/ (some correct, some distorted or substituted) and asked listeners to rate consonant accuracy by reflecting their judgements on a computerized version of a VAS where the target phonemes were represented at either end. A click at either extreme of the scale represented a highly accurate exemplar of either phoneme, and a click closer to the center represented a distortion, or "intermediate form". Results revealed that listeners were able to perceive a small, but statistically significant, difference in phoneme productions that was not represented in broad phonetic transcription. As such, these authors concluded that this result revealed a covert contrast in the child's speech.

Statement of Purpose and Research Questions

Previous research has established that children with CIs experience a large amount of variation in speech production ability during their development (Blamey et al., 2001; Ertmer & Goffman, 2011; Flipsen, 2011; Sarant et al., 2001; Spencer & Guo, 2013) that can be partially explained by differences in neural plasticity among individuals (Lee et al., 2001; Sharma et al., 2002; Wilson & Dorman, 2008) and acoustic cue use (Giezen et al., 2010; Moberly, et al., 2014). Other investegators have identified and classified speech sound errors, created phonetic inventories and denoted change over time in the accuracy of phoneme production by children with CIs using both broad and narrow transcription (Blamey et al., 2001; Chin, 2003; Ertmer et al., 2012; Flipsen, 2011; Spencer & Guo, 2013). They found that children with CIs develop speech similarly to children with NH, with a few exceptions.However, it is not clear why some phonemes developed in non-typical ways, while others resembled what is seen in children with NH. One possibility is that children with CIs either do not have adequate access to certain parts of the speech signal, or they are attending selectively to inappropriate acoustic cues.

Examination of a phenomenon that has been shown to be common among typically developing and disorder children, covert contrast, may shed light on this question. The existence of covert contrast in children with CIs has not yet been etablished, but may have important clinical implications regarding therapy progress. It is possible that using broad transcription, coupled with a measurement tool that is sensitive to subtle changes in phoneme productions would demonstrate covert contrast in this population. Since /t/ was repeatedly shown to be unusually late-developing in children with CIs (Blamey et al., 2001; Ertmer et al., 2012; Spencer & Guo., 2013), it was chosen as the phoneme of interest in this investigation. Specifically, this study seeks to answer the following questions:

1. Do children with CIs produce /t/ as accurately as children with NH who have similar gross articulatory ability?
2. When children with CIs and NH substitute another sound for /t/, is there a significant perceptible difference between the substitution and the correct production of the other sound?

Chapter 2

Methods

Participants

Speakers. Two groups of preschool-aged children participated in this study: children who used CIs (Experimental Group) and speech-age matched peers (Control Group). All of the children were recruited as part of a larger study that examined the influence of speech production abilities on speech perception scores of children with CIs (Gonzalez, 2013). Parents of the participants provided the original investigators with detailed demographic information via questionnaire, which allowed them to rule out several exclusionary characteristics. These included: cognitive delay or impairment, cognitive or psychiatric disabilities, and primary language use other than English.

Experimental group. The experimental group included nine congenitally deafened children with profound sensorineural hearing loss (5 females, 4 males) who had been fitted with CIs (see Appendix A for additional demographic information). All participants in the CI group: 1) were implanted by 3 years of age, 2) had at least 12 months of CI device experience at the time of testing, and 3) used an oral mode of communication exclusively prior to implantation. This was important because children who use oral communication are encouraged to communicate using speech. Previous research has shown that children trained in the oral tradition have superior consonant acquisition (Conner, Hieber, Arts, & Zwolan, 2000).

Control group. Members of the control group were selected from a pool of 24 possible participants. Inclusion criteria were as follows: 1) between the ages of 3-5 years, 2) normal hearing (i.e., hearing thresholds ≤ 20 dB HL from 250 Hz to 4000 Hz), and 3) no middle ear involvement at the time of testing. Of the 24 children whose parents had consented for their children to participate in this study, eight were determined to have appropriate speech production abilities to serve as matches to the experimental group. The control group participants (5 females, 3 males) were between the ages 2:8 to 5:1 years (M=4:0).

Each child with a CI was matched to a child with NH by articulation ability using scores from a standardized test of articulation and gender, when possible. Raw scores for each participant (the sum of all articulation errors) were converted into standard scores based on hearing age for the experimental group and chronological age for the control group. Hearing age was defined as time since device activation. Participants were considered "matched" if their respective standard scores fell within the 95% confidence interval of a child with NH (see Appendix B). One matched pair (CI06 and NH17) did not meet this criterion; the standard score for the child with a CI was higher than the NH child based on hearing age, and their 95% confidence intervals did not overlap. However, the two children were exactly the same age (56 months), were both female, and received similar raw scores (CI=3, NH=6). Given these circumstances, the two were considered to have similar articulation abilities and were paired.

Listeners. Thirty-three graduate students in speech-language pathology were recruited to participate as listeners in this project. They had completed a phonetics course, voluntarily participated in the listening experiment, and received no compensation. No other demographic data were collected on the listeners.

Materials

Speech and language data were obtained from the previous investigation (Gonzalez, 2013). These tests include the *Peabody Picture Vocabulary Test 4* (PPVT-4; Dunn & Dunn, 2007), the OlimSpac (Boothroyd et al., 2006), and the *Goldman-Fristoe Test of Articulation 2* (GFTA-2; Goldman & Fristoe, 2000). These tests will be described briefly below.

Peabody Picture Vocabulary Test 4. The PPVT-4 (Dunn & Dunn, 2007) is a standardized test of receptive vocabulary. During administration, a word was spoken aloud by the examiner and the child was asked to make a selection by pointing to the corresponding picture in a field of 4 possible choices. All subjects achieved a score that was within one standard deviation of the mean score for their typically developing age-matched peers.

Group matches were not determined using this measure. Rather, it was deemed an appropriate tool to assess differences in vocabulary acquisition that may have impacted phonological representations, and thus speech production (Gonzalez, 2013). Since all children scored within one standard deviation of the mean, receptive vocabulary was effectively ruled out as a confounding factor to differences in speech production ability.

OlimSpac. The OlimSpac (Boothroyd et al., 2006) is computerized software program that was originally designed to clinically assess the ability of children with hearing loss to perceive six phonologically significant speech contrasts (see Table 1).

During testing, each subject was presented with pre-recorded VCV nonwords over a loudspeaker, while seated in front of a computer monitor in a sound-proof booth. The children were instructed to “watch the screen”, listen for each sound presentation, and repeat the nonsense word to the best of their ability. Each OlimSpac stimulus item was presented in both an auditory-only, and auditory-visual condition. During the auditory-only trials, the screen displayed a colorful image that changed color when the stimulus played. During the auditory-visual trials, the screen displayed an adult female's face as she pronounced the stimulus accurately.

Each speech contrast was represented at least twice by different phonemes. For example, perception of consonant continuance was assessed with both a t/s contrast, and a □/□ contrast. Selected contrasts were consistent among subjects, but presented in a random order during each test session. Cumulatively, each child imitated 16 VCV nonwords in each condition (auditory-visual, and auditory-only), for a total of 32 imitated productions per child. Their imitated productions were recorded for future analysis using an Olympus ME52 directional lapel microphone connected to an RCA VR 5220 digital voice recorder.

Table 1. OlimSpac Speech Contrasts

Speech Contrast	Example
Vowel height	/udu/ vs. /ada/
Vowel place	/utu/ vs. /iti/
Consonant voicing	/ata/ vs. /ada/
Consonant continuance	/iti/ vs. /isi/
Pre-alveolar consonant place	/upu/ vs. /utu/
Post-alveolar consonant place	/utu/ vs. /u□u/

Goldman-Fristoe Test of Articulation-2. The GFTA-2 (Goldman & Fristoe, 2000) is a standardized test of articulation, which is considered to be the gold-standard assessment of speech production skill among speech-language pathologists. This test is typically used to assess the acquisition and accuracy of 39 English consonants and consonant clusters at the single-word level in all three word positions (initial, medial and final). During testing, the examiner shows a colorful picture to the child in order to elicit a word containing one or several target consonants. If the child does not know the word, or produces an alternate word, the examiner provides the correct target, which the child imitates after a delay.

For this project, a graduate student in speech-language pathology (SLP) phonetically transcribed subject responses, which were then reviewed by a second graduate SLP student. A third “expert” clinician, who was a certified SLP, was consulted to resolve discrepant transcriptions. Once the transcriptions were verified, noted consonant and consonant cluster errors were added up to determine a raw score. Each raw score was converted to a standard score and 95% confidence interval. For the NH group, standard score conversions were based on chronological age. Standard scores for the CI group, however, were calculated using the subjects’ “hearing age.”

Development of Experimental Material

In order to study covert contrast, one has to have multiple productions/misarticulations of the target phoneme. Given the stimuli available to this project, the multiple productions of /t/ from the OlimSpac were selected for further perceptual analysis. It is recognized that each child's production of a phoneme on the OlimSpac is essentially an imitation of what they thought they heard, which is different than asking a child to name an object/picture with the desired target phoneme. However, given the similar vocabulary levels and speech production skill, as well as the high accuracy levels obtained on subtests of the OlimSpac (see Appendix C) for all participants, it seems reasonable to assume that any productions taken from the OlimSpac will be representative of the child's ability to produce the target sound and not significantly impacted by misperception of the target.

Selection of consonants. OlimSpac recordings from the nine members of the experimental group, and the eight articulation-matched control subjects were analyzed for accuracy by the current investigator, who did not participate in the testing of the participants or scoring of the GFTA-2. OlimSpac productions were transcribed and determined to be either a correct production or a clear substitution. Distortions were counted as correct, but were anomalous for some reason (inappropriate aspiration, imprecise production, resembled a similar phoneme, etc). A confusion matrix was then generated based on these categories, as illustrated in Table 2.

Table 2. CI user's productions, in percentage, for each error type.*

	Child's Production									
		/p/	/b/	/d/	/t/	/s/	/⍰/	/⍰/	Other	Dist*
OlimSpac Target	/p/	72.22	2.78	8.33	8.33				8.33	19.23
	/b/	13.89	63.89	11.11	8.33				2.78	13.04
	/d/			62.16	27.03				10.81	17.39
	/t/		1.41	7.04	73.24	1.41	2.82	14.08		34.62
	/s/	2.78		2.78	2.78	52.78	22.22		16.67	63.16
	/⍰/			5.56	5.56	2.78	66.67	13.89	5.56	12.50
	/⍰/			2.78	19.44		22.22	55.56		15.00

**Percentage of "correct" productions that were less-than-perfect. See above for examples.*

Phoneme substitutions for /t/ (the phoneme of interest in this investigation) were selected using the data provided by the confusion matrix. When members of the CI group substituted another phoneme for /t/, they were most likely to produce /□/ or /d/. These productions accounted for 14.08% and 7.04% of /t/ errors, respectively. When /□/ and /d/ were produced incorrectly, they were similarly very likely to be produced as /t/ (see Table 2). Therefore, /t/, /d/, and /□/ were selected as the phoneme productions of interest in this experiment. The result was four error categories (i.e., transcription categories) for two possible substitutions for /t/: t/t, t/d, d/t, and d/d or t/t. t/□, □/t and □/□.

The investigator was interested in studying covert contrast, as previously discussed, which can be found in the "intermediate forms" of children's speech as they are developing consonants. The selection of phonemes that were observed to be frequently substituted for each other increased the likelihood of discovering covert contrast.

These phoneme choices were particularly appropriate because one differed in voicing (t/d) and the other in manner of articulation (t/□). Place of articulation consistently has been shown to be poorly transmitted by CIs (Clark, 2003; Giezen et al., 2010; Pisoni et al. 1999), so a place contrast (such as t/p) was not included in this experiment. In addition, since coronal place of articulation has been shown to be well-transmitted by the speech processors of CIs, one can assume that the speakers in this study received as much acoustic information as possible from their speech processors for adequate /t,d,□/ perception (Dillon, Cleary, Pisoni, & Carter, 2004).

Isolation of VCV nonwords. After the appropriate VCV nonwords were selected, the speech samples were digitized at 20,000 Hz using Praat (Boersma & Weenink, 2013). For each subject, every opportunity of the three target consonants was isolated from the recordings of the participants and saved as a separate .wav sound file. Each child had eight opportunities to produce /t/, and four opportunities each to produce both /d/ and /□/. All target stimuli, with vowel contexts included, are listed in Table 3. No effort was made to control for listening condition because the original investigators found no significant difference in consonant accuracy between the auditory-only, and the auditory-visual conditions for either the NH or CI group. This decision also increased the number of nonwords and consonant errors available for the test procedures.

Table 3. OlimSpac stimuli that met the selection criteria

Isolated OlimSpac Stimuli
/ata/
/utu/
/iti/
/ada/
/udu/
/i□i/
/u□u/

Isolated OlimSpac samples for each subject were screened for clarity, and poor samples were removed from the experimental pool. The poor samples were deemed unusable because of interfering noise, low intensity, the presence of other voices, or the subject's lack of cooperation. This led some transcription categories to be unevenly represented in the final group of stimuli. However, care was taken to ensure that each subject and each phoneme was represented as proportionately as possible (see Tables 4 & 5).

Table 4. Number of VCV non-words for each transcription category in block 1.

	Correct [t]	[d] : [t]	[t] : [d]	Correct [d]
CI	9	5	7	5
NH	11	3	5	5

Table 5. Number of VCV non-words for each transcription category in block 2.*

	Correct [t]	[□] : [t]	[t] : [□]	Correct [□]
CI	10	8	6	4
NH	10	1	5	6

* *Note that there was only one occurrence of a /□/:/t/ substitution in NH children.*

Acoustic equalization of sound files. After the experimental pool of isolated speech samples was assembled, the files were individually “cleaned” of noise using the noise reduction feature of the audio editing software Audacity® (SourceForge, 2013). To do this, a segment of silence was selected from the audio track. “Silence” is designed to measure the ambient noise on a tape, not related to the experimental stimulus. The software then attenuated those "noise" frequencies that had been identified.

After all files had been processed in this way, they were then collectively normalized. This was done to reduce the variation in loudness between samples. Some children spoke with a significantly reduced vocal intensity, and some children’s mouths were farther away from the lapel microphone than others. This created variability in loudness among the raw files. In order to achieve a uniform loudness, a single file was selected that had 1) an appropriate loudness level and 2) a clear-sounding, noise-free signal. This file was used as a benchmark and all other files were normalized relative to it. The Audacity® (SourceForge, 2013) software automatically performed these alterations after the baseline file was selected, and all other experimental files were input into the program.

Selection of measurement tool. Previous research using rating scales to assess the presence of covert contrast utilized VAS successfully (Munson et al., 2010; Munson et al., 2012; Schellinger et al., 2010). The VAS has been shown to be both reliable and valid when used to measure other subjective qualities that cannot otherwise be directly measured, such as vocal quality and pain. In like fashion, equal-appearing interval (EAI) scales have been shown to be useful in rating subjective characteristics. According to Yiu & Ng (2004), there is a moderate correlation (.56-.76) between VAS and EAI scale ratings for identical stiumli. However, the EAI ratings showed significanlty higher intra-rater reliability (VAS agreement = 0.57; EAI agreement = 0.73). This may make the EAI scale more appropriate for use in a clinical setting where intrarater reliability is of utmost importance. When assessing a child's progress toward a target sound, it is essential that the clinician's judgements are consistent. Therefore, the use of an EAI scale, rather than a VAS was selected for this experiment.

Programming the experiment. All altered sound files were divided into two blocks. Target [d] and [□] productions were separated into different blocks: all samples of target [d] were placed into block 1, and all samples of target [□] placed into block 2. Samples of target [t] were equally distributed between the two blocks. The distribution of subjects and target sound was carefully checked to ensure proportional assignment between the two groups. After this assignment was completed, each block contained 60 unique speech production samples. The sound files were randomized using a random number generator, and programmed in a fixed order into the block. File order was then quasi-randomized to ensure that no more than two similar-sounding files, (either by stimulus or subject) were presented consecutively. The first 12 files presented in each block were duplicated for presentation at the end of the block in order to assess intra-rater reliability. The files were presented in the same fixed randomized order each time the experiment was presented.

The ECoS experimental design software (Avaaz, 2002) allowed the creation of “blocks” as a means to separate the experiments. Each experiment block was preceded by a training block consisting of 10 novel sound files that were not utilized in the experiment blocks. These files were carefully selected from the original pool of possible experimental stimuli. The training stimuli were not randomly selected, but were chosen to demonstrate a range of different variables in which the experimental stimuli would differ. They were meant to provide the listeners with a representative sample of sound quality, vocal quality, vocal intensity, and consonant clarity or ambiguity, in order to allow them to become familiar with the rating task.

Equipment

The experiment file and all associated sound files were placed on two identical Dell Optiplex desktop PCs with identical internal hardware. The experimental sounds were directed through Califone circumaural headphones connected to the PC through a 3.5mm headphone jack. Listeners were tested in a quiet environment, and in tandem, when possible.

Procedures

When the listeners arrived on the day they were to participate, they were asked to fill out a brief questionnaire with no identifying information in order to ensure consistency in listener characteristics. All listeners self-reported: adequate hearing, typical neurological status and cognition, and English as a first language. Additionally, no listener showed evidence of a speech or language disorder, as judged by the examiner.

The listeners were then given instructions regarding the experimental task. They were told that they would be listening to children producing VCV nonwords, and were given an example (like [ada]). They were informed that some children would sound very clear and others would be harder to understand. They were also told that some trials would be shorter than others, some would have different intonation, and some would be more distorted than others. They were asked to not focus on these differences but to attend to the adequacy of the sound production. Other instructions included the following:

- Don’t spend too much time on finding the “right” answer. It is best to go with your gut.
- Focus on the consonant in the middle of the vowel. Try to ignore everything else.
- You may repeat the trial as many times as you wish.
- You may adjust the volume to a comfortable listening level.

The listeners then were shown a duplicate of the equal appearing interval (EAI) scale that would be presented during the experiment. The EAI was a linear scale with 7 possible selections (see Figure 1). The listeners were directed to click a point on the scale that most closely corresponded to their interpretation of the phonetic accuracy of the consonant presented in each trial. A score close to either extreme of the EAI indicated a very accurate production of the nearest phone, with 1 or 7 being a “perfect” production of that phone. A score of 4 would represent an inability to distinguish between the two phonemes (a perfect blend). In block 1, listeners rated the subjects' attempts at /t/ and /d/. In block 2, they rated the subjects' attempts at producing /t/ and /□/.

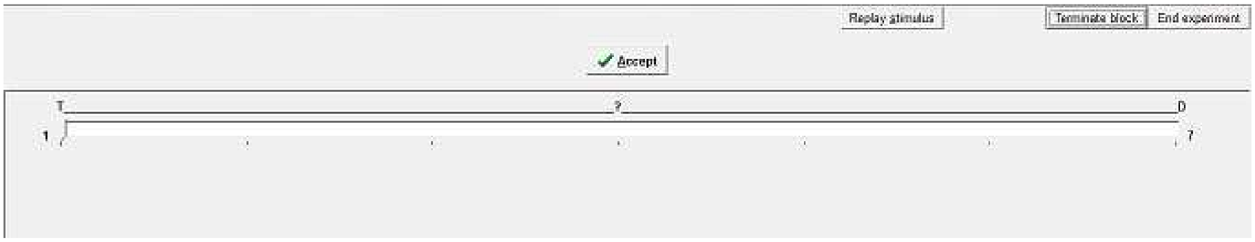

Figure 1. A reproduction of the EAI seen by listeners during block 1

They were told that during the experiment, they would rate the "phonetic accuracy" of the consonant contained within each VCV nonword on the scale provided, and were encouraged to use the entire EAI scale when rating the sounds. Finally, they were informed that they would have an opportunity to "practice" using the scale during the short training blocks before beginning the longer listening experiment.

When it was indicated that the verbal instructions were understood, the listener was seated in front of a computer monitor with the instructions "Click OK or press the Enter key to start the experiment" displayed on the screen. After the listener initiated the experimental block, written instructions were presented (see Appendix D). The listener again selected "OK" and the first training stimulus was automatically presented. After making a selection on the EAI to indicate the quality of the consonant presented, the listener clicked "Accept" and the next stimulus automatically was presented. A "Replay" button was visible during each trial and the participant was allowed to replay the stimuli as many times as needed. Once a selection was made, and the listener clicked "Accept", the rating could not be altered. After the 10 training trials were rated, a second set of written instructions was displayed (see Appendix D). The listener clicked through as before, and listened to the each of the 72 experimental stimuli. The second training block and experimental block were presented in an identical way. After all sounds were rated, a box appeared, notifying the listener that the experiment had ended.

Data analysis

Intra-rater reliability was calculated using the duplicated trials from each listening task. Listener responses were separated by transcription category (TC) (correct [t], substituted [t], substituted other, correct other) and then an average for each TC was computed by participant. A 3-way repeated measures ANOVA was performed to analyze listener response data. The independent variables were group (NH or CI), TC (error type), and contrastive choice (CC;. whether the comparison involved t/d or t/□). . The dependent variable was the average perceptual score. Post-hoc analyses were conducted as needed and effect sizes were calculated.

Chapter 3

Results

The primary purpose of this investigation was to compare the accuracy of consonants produced by children with NH and children with CIs who were matched for articulatory ability. The author sought to 1) examine the accuracy of /t/ productions in children with CIs, which had been shown to be unusually late-developing in several previous studies (Blamey et al., 2001; Chin, 2003; Ertmer, et al., 2012), and 2) quantify subtle phonetic differences in correctly produced consonants and substituted consonants that a less sensitive articulation measure could not detect. To this end, listeners rated productions of /t/ and the phonemes that were found to be most often substituted for it, /d/ and /□/. Consonant ratings were made on a 7 point EAI scale. A rating of 1 represented a “perfect” production of /t/, and 7, a “perfect” production of the other consonant: /d/ (block 1), or /□/ (block 2). The center value on the scale, 4, signified an intermediate production, or a “perfect blend” of the two consonants. It was anticipated that the use of EAI scale would more accurately portray listeners’ perception of consonant accuracy, and reveal phonetic differences between productions if they existed. Specifically, this study aimed to answer the following questions:

1. Do children with CIs produce /t/ as accurately as children with NH who have similar articulatory ability?

2. When children with CIs and NH substitute another sound for /t/, is there a significant perceptible difference between the substitution and the correct production of the other sound?

Intra-Rater Reliability

First, an examination of intra-rater reliability was undertaken to determine whether or not listeners were using consistent criteria to judge the speech stimuli. A common method of assessing intra-rater reliability is to calculate the percentage of repeated ratings that are within ± one scale value of the original rating (Kreiman, Gerratt, Kempster, Erman, & Berke, 1993). Over both experimental blocks, each listener rated 24 stimuli twice. For each listener, the percentage of responses to duplicated stimuli that were within ± one scale value of the original rating was calculated, and these values were averaged across listeners. Calculations revealed that overall, 88.1% of duplicated trials were within ± one scale value of the original rating. Of these, 56.6% were in exact agreement. Hence, listener judgment reliability was determined to be good.

After participation, listeners were asked to provide feedback on their performance. The majority of listeners anecdotally reported that the sounds presented in the second block were harder to differentiate. One listener noted that the sounds presented in block 2 seemed to be "less precise". Another said that they were simply "harder to distinguish". This observation was borne out by the reliability data. Intra-rater reliabilities for both blocks were independently sufficient (>80%), but listeners were consistently less reliable in their responses for block 2. Detailed data regarding reliability for each block can be found in Table 6.

Table 6. Percentage of listener ratings that agreed between duplicated trials.

	Mean Exact Agreement	Mean ± 1 scale value (SV)	Mode ± 1SV	Range ± 1SV
Block 1	61.7%	89.5%	100%	66.7 – 100%
Block 2	51.5%	86.7%	83.3%	58.3 - 100%

Selection of Variables

Independent variables included: experimental group (CI vs. NH), TC (4 levels of correct and substituted productions) and CC (/d/ or /□/). The principal investigator's determination of "correct" or "substituted" for each VCV trial was used to assign them into a TC (see Chapter 2). All samples of VCV non-words were grouped into their assigned transcription categories, and then ratings for each transcription category were averaged for each listener.

Results of the Statistical Analyses

A 3-way repeated measures ANOVA was conducted to analyze the influence of group (i.e., hearing status), TC and CC on perceptual ratings. Results revealed a significant 3-way interaction. However, differences across the experimental blocks were not of primary interest, so they will not be discussed further. Statistical analysis further revealed that two of the three main effects were significant, experimental group, $F(1,32) = 27.99$, $p < .001$, $\eta_p^2 = 0.467$ and transcription category (TC), $F(3,96) = 760.70$, $p < .001$ $\eta_p^2 = 0.960$. These results suggest that differences were evidenced across groups an TC. However, there was also a significant interaction between group and TC, therefore, the research questions will be addressed within this interaction. The main effects and interactions are summarized in Table 7.

Table 7. Summary of results the 3-way repeated measures Analysis of Variance.

Main Effect	**df**	**Mean Square Error**	***F***	***p***	η_p^2
Group	1,32	7.548	27.99	<.001	.467
CC	1,32	1,314	1.141	.293	.034
TC	3,96	272.597	760.70	<.001	.960
Group x CC	1,32	4.325	18.70	<.001	.369
Group x TC	3,96	6.699	25.56	<.001	.444
CC x TC	3,96	48.735	149.74	<.001	.824
CC x Group x TC	3,96	49.044	137.33	<.001	.811

**CC= contrastive choice; TC= transcription category*

Accuracy of /t/ productions. In the past, /t/ has repeatedly been shown to be unusually late developing in children with CIs. The first goal of this project was to confirm this observation by examining listener perceptions of accuracy. This was best satisfied by examination of the significant Group x TC interaction, $F(3,96) = 25.562$, $p < .001$, $\eta_p^2 = 0.444$. This finding suggests that differences in transcription category were dependent upon group. Post-hoc testing results using paired samples t-tests with a Bonferroni correction ($p = .004$) revealed that 3/8 paired comparisons of interest were not significant: t/t in both blocks, and □/[□ (see Figures 2 and 3). In other words, [t] and [□] productions were not judged to be significantly different across groups; however [d] productions were significantly different across groups. Therefore, when the author judged the production to be a /t/, children with CIs successfully produced /t/ as accurately as their normal hearing peers.

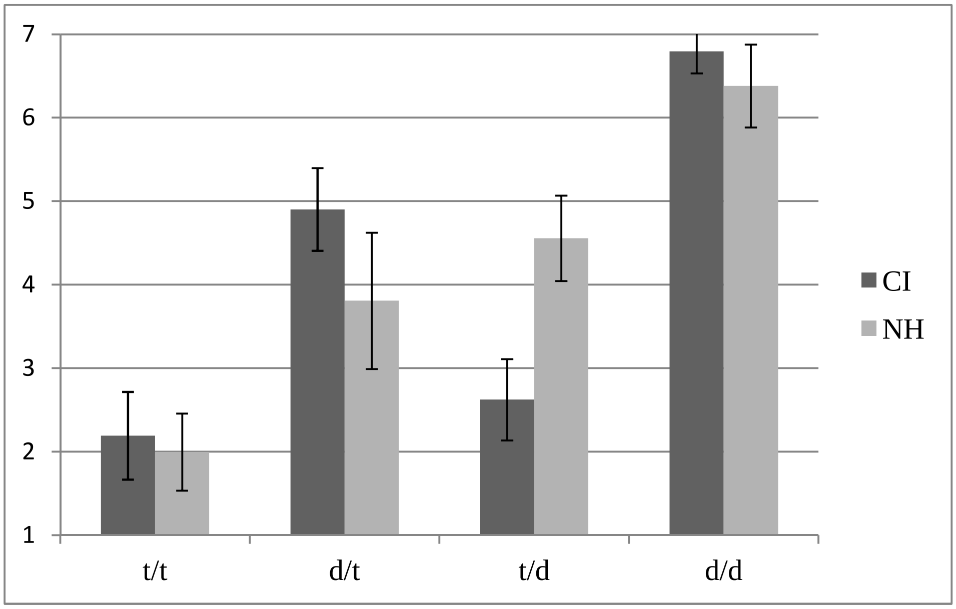

Figure 2: Differences in listener perceptions of consonant accuracy for block 1

**Covert contrast is shown in the d/t and t/d contrasts.*

While the above findings demonstrated no group differences for /t/, it did not address the issue of whether or not /t/ was produced in error by children with CI more often than other phonemes or when compared to productions from NH children. Overall error frequency taken from OlimSpac testing was determined to provide enough additional relevant information to warrant analysis.

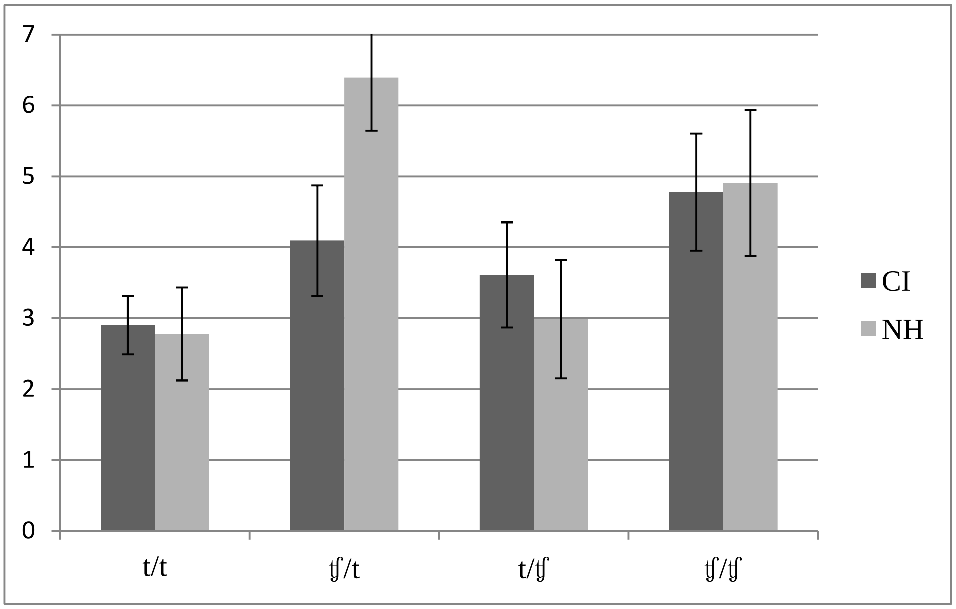

Figure 3. Differences in listener perceptions of consonant accuracy for Block 2

**Covert contrast is shown in the [□]:[t] and [t]:[□] contrasts.*

A confusion matrix of CI group productions had previously been generated when selecting contrastive consonant choices (as described in Chapter 2). To compare error frequencies between groups, a second confusion matrix (of NH productions) was created. The combined results are found in Table 8. Examination of this confusion matrix revealed that NH children produced /t/ accurately in 87.93% of opportunities, whereas children with CIs produced it accurately in 73.24% of opportunities. In addition, this discrepancy between the NH and CI groups in phoneme accuracy was found to be proportional across all consonants attempted during OlimSpac testing. In other words, the children with CIs were uniformly less accurate. To be specific, [t] was found to be perceptually less accurate when all attempts were considered, but no more so than any other consonant presented.

Table 8. A Confusion matrix that compares production accuracy and error patterns across NH and CI groups.

	Child's Production- **CI** **NH**								
		/p/	**/b/**	**/d/**	**/t/**	**/s/**	**/□/**	**/□/**	**Other***
OlimSpac Target	**/p/**	**72.22** **81.25**	**2.78** **15.63**	**8.33**	**8.33**				**8.33** **3.13**
	/b/	**13.89** **9.38**	**63.89** **78.13**	**11.11** **6.25**	**8.33** **3.13**				**2.78** **3.13**
	/d/		**3.13**	**62.16** **81.25**	**27.03** **12.5**				**10.81** **3.13**
	/t/		**1.41**	**7.04** **8.62**	**73.24** **87.93**	**1.41**	**2.82**	**14.08** **3.45**	
	/s/	**2.78**		**2.78**	**2.78**	**52.78** **68.75**	**22.22** **9.38**		**16.67** **21.88**
	/□/			**5.56**	**5.56**	**2.78** **21.88**	**66.67** **71.88**	**13.89** **3.13**	**5.56** **3.13**
	/□/			**2.78**	**19.44** **21.88**		**22.22** **6.25**	**55.56** **71.88**	

*"Other" refers to phoneme productions not tested on the OlimSpac.

Perceptible contrast between substitutions and correct targets. The second purpose of this investigation was to determine whether or not covert contrast was present in the speech of children with CIs, and if so, within what phonemes. This question was also best addressed by examination of the significant within group post hoc results of the Group x TC interaction. All within group paired comparisons for both the CI and NH group were significant. In other words, the "correct" /t/ was significantly different from the /t/ used as a substitution, as well as when a /d/ or /□/ was used for a /t/. In contract, post hoc testing revealed significant differences in the similarity of /t, d, □/ productions across groups when they were used as substitutions for other phonemes (i.e., t/d, d/t, t/□, □/t). These findings illustrated covert contrast. In addition, there were significant differences in the patterns of covert contrast across groups. As illustrated in

Figures 2 and 3, all four paired comparisons involving phoneme substitutions across groups were significant (p < .001). When children with CIs substituted d/t, it was more [d]-like and when they substituted t/d, it was more [t]-like. The opposite pattern was noted in NH children. However, a different pattern was noted for /□/. For children with CIs, the t/□ substitution was more /□/-like than for NH hearing children. While there was a significant group difference for the □/t substitution, there was only one instance of this error in the NH group, so a group comparison is not appropriate. Nevertheless, the [t] production here for the children with CIs was more[□]-like.

Summary of Results

The purpose of this project was to determine whether the development of /t/ in children with CIs differed from their articulation-matched, NH peers, and to explore listener perceptions of that consonant when correct, and when in error. The first stated purpose regarding /t/ accuracy was addressed by comparing perceptual ratings of correct [t] for both groups. As there was no difference found, CI productions of /t/ were determined to be as accurate as those of their NH peers when they were produced successfully. However, this result represented success across participant groups. It did not compare /t/ accuracy to other phonemes attempted during the administration of the OlimSpac. When a confusion matrix was generated to depict OlimSpac performance, the NH group was noted to outperform the CI group across all phonemes. This would suggest that /t/ was not uniquely poorer in the CI group, but instead the children with CI evidenced poorer phoneme accuracy in general.

The second aim of this study was to assess whether phoneme substitutions were perceptibly contrastive to correct productions of the substituted phoneme and if these differences were similar across groups. This is commonly referred to as “covert contrast” in the literature. To this end, listener’s ratings of substituted [t], [d] and [□] productions were compared to correct tokens of their most common substitutions. Covert contrast was said to be present when the substitutions were more /t/-like than the correct productions. It was determined that a statistically significant contrast did exist between correct [d] and [□], and those phonemes when substituted for [t] for both groups. A significant contrast was observed for both groups between correct [t] and t/d substitutions; however the nature of this contrast differed by group. In the second block, listeners perceived a difference between correct [t] and t/□ substitutions in the CI group productions. This error pattern did not occur often enough in the NH group to be reliably analyzed. .

Chapter 4

Discussion

Other investigations of speech development in children with CIs have shown initial accelerated growth, followed by a plateau where consonant order of acquisition generally mirrors that of NH children, but is slower (Blamey et al., 2001; Serry & Blamey, 1999; Spencer & Guo, 2013). A notable exception to this pattern, /t/, has been shown to be acquired later-than normal in several investigations (Blamey et al., 2001; Chin, 2003; Ertmer et al., 2012). The primary purpose of this project was to confirm that /t/ is a phoneme that is more difficult for children with CIs to develop. The secondary purpose was to assess the development of /t/ by looking for differences in correct and incorrect /t/ productions. This was achieved by examination of listener ratings of CI and NH productions of [t] and its most common substitutions, [d] and [□].

A three-way ANOVA was performed to determine the differences in perceived consonant accuracy across three independent variables: group, TC, and CC. Results revealed a significant main effect for both group and TC, as well as a significant Group x TC interaction. These results indicated that children with CIs did not show an unusually delayed development of /t/ when produced correctly. However, differences between groups were apparent in substitutions of [t] for target /d/ and /□/ productions. This discussion will begin by first presenting the results of the experimental findings relative to the two questions the present study sought to address. Next, strengths and limitations of the present study will be presented. Finally, clinical implications and future research needs will be explored

The Development of /t/ in Children with CIs

The first research question was addressed in two ways. First, listeners' perception of correctly produced /t/ was compared across groups. A three-way repeated measures ANOVA found that listeners did not perceive a significant difference between groups when only correct /t/ productions were considered. Hence, the /t/ produced by children with CIs was no less accurate than when produced by NH children when speakers were matched for articulation ability.

Lack of a significant finding may be related to advances in CI speech processing technology, as several studies showing delayed /t/ acquisition were more than 10 years old (Blamey et al., 2001; Chin, 2003). Another possible explanation is differing years of device experience in the children from different studies. The two other studies that previously showed delayed /t/ development recruited subjects that had 2 years, and 0-3 years of device experience respectively (Ertmer et al., 2012; Spencer & Guo, 2013), whereas the children in the present study had an average of 3 years device experience. Given that children with CIs gain speech sound accuracy rapidly at first, and then slowly thereafter, it is possible that our subjects had reached the "plateau" stage of development, whereas subjects in the comparison studies were still experiencing rapid growth. If this were the case, our subjects' accuracy for all consonants, including /t/, would be understandably higher.

Accuracy of /t/ was further analyzed by examining the overall accuracy of all phoneme attempts during OlimSpac testing. This was accomplished using a confusion matrix that illustrated error frequencies of all OlimSpac stimuli for both groups. This analysis revealed that children with CIs erred more often than children with NH on all phonemes, including /t/. However, the error pattern was proportional; /t/ did not stand out as unusually delayed. If /t/ development were truly anomalous in the present study, one would expect the discrepancy between CI and NH percentage of tokens produced correctly to be much larger for /t/ than other phonemes. However, this was not the case. Further, it was found that both groups produced /t/ accurately more often than any of the other six phonemes considered on the OlimSpac (/p,b,d,s,□,□/). This pattern is also inconsistent with the findings of Ertmer et al. (2012) who reported that children with CIs produced initial /t/ less accurately than both /d/ and /□/. Thus, previous findings of unusually delayed development of /t/ in children with CIs were not confirmed in the present study.

These differences may also be attributable to differences between previous studies in the demographic characteristics of their experimental groups. The children with CIs who participated in this study came from families with higher parental education and relatively high socio-economic status (SES; see Appendix A). It is well established that children with a higher SES tend to experience better language-related skills, such as vocabulary development, expressive and receptive language, and phonological awareness (Raviv, Kessenich, & Morrison, 2004). Children who have generally superior language faculties will naturally outperform other children on any language-related task, including speech production, regardless of hearing status.

Covert Contrast

The second purpose of this study was to establish and present evidence for the occurrence of covert contrast in children with CIs. It was hypothesized that listeners would perceive covert contrast in the speech of the CI group because it has been found consistently in children's speech (both typical and disordered) as they gradually acquire consonant contrast (Schellinger et al., 2010). To address this question, listener ratings of substituted tokens of target /t, d, □/, and ratings of correct tokens of the same phonemes, were compared. Covert contrast was said to have been present if the two sounds (one substituted, and one correct) were transcribed as identical but were perceived by listeners to be significantly different. This type of perception is related to a gradual progression from a pure substitution to a more adult-like target. As a child experiences this progression, substitutions will sound less “pure” and more like their intended target. However, a child who had no phonological representation of the phoneme would produce a substitution that was indistinguishable from the contrastive target.

Post-hoc comparisons of the group by TC interaction showed that the both the CI and NH groups demonstrated significantly perceptible differences between correct and substituted phoneme productions for all phonemes of interest. The voicing contrast was more evident in the listener's perceptions of the error phonemes for children with NH than the children with CIs, demonstrating that NH children who produce voicing substitutions are closer to mastering an adult-like stop voicing contrast than children with CIs. On the other hand, children with CIs appeared to struggle with the voicing distinction. That is, t/d substitutions sounded more like [t] than d/t substitutions sounded like [d]. This confirms Gonzalez et al.’s (2013) observation that the children with CIs struggled with either the perception or production of voicing more than other contrasts. Given that consonants within the OlimSpac stimuli are sandwiched between vowels, one would expect any speaker to struggle with the more complex perception of a “voiced-voiceless-voiced” syllable (as in /ata/), than with a voiced-voiced-voiced syllable (as in ada). That the opposite trend was observed in children with CIs was surprising.

A comparison could not be made across groups for the accuracy of [t] in block 2 due to the small number of [□] : [t] errors produced by the NH participants. However, sufficient CI group productions were represented for both error types, allowing comparison in that group. Results revealed that, as in block 1, listeners perceived a difference between correct and substituted phonemes in that block 2. However, unlike the voicing contrast pattern seen in block 1, the manner contrast in block 2 was perceived more as something between /t/ and /□/. These errors may be related to the poor perception of the spectral cues associated with /t/ and /□/ (Hicks & Ohde, 2005), given the imitative nature of the OlimSpac task. Nevertheless, evidence of covert contrast in both substitution options for the /t/ phoneme would suggest that children with CIs have developed a good phonological representation for /t/, and that their substituted productions are contrastive despite possible perceptual ambiguity. The demonstrable presence of covert contrast in children with CIs reiterates the conclusion that this group has acquired /t/, even if they have not completely mastered it.

More interesting is the finding that the children with CIs produced a perceptible contrast between correct [t] and t/□ substitutions, whereas the children with NH did not. While it was expected that children with NH, who were only 4;0 years on average, would not have a mature phonological representation for the later-developing [□] phoneme, it was surprising that the children with CIs did. The CI group participants had an average chronological age of 4;9 years, which is approximately the age at which Smit et al. (1990) reported that typically developing children acquired /□/. However, the CI group only had an average of 3 years,2 months of robust hearing experience. Perhaps some feature is present in the speech of children with CIs that causes [t] substitutions to sound more [t]-like when contrasted with /d/, and more [□]-like when contrasted with /□/. It is possible that these children with CIs are attending to different acoustic cues, or find certain cues more salient than children with normal hearing. Idiosyncratic attention to an acoustic cue that is more readily transmitted by the CI, for example aspiration, may explain these patterns. However, acoustic analysis is needed to test this hypothesis.

More than simply demonstrating the accuracy or emergence of a certain phoneme, the presence of covert contrast uncovered in this investigation speaks to the phonological development of children with cochlear implants. Logically, in order to produce any contrast, whether covert or overt, a speaker must first *perceive* the contrast. If the children with CIs perceived no difference between /t/ and /d/, or /t/ and /□/, then /t/ would be produced identically to the latter two phonemes. However, since they do produce a contrast, that they are receiving enough auditory information to develop an accurate underlying representation of /t/, and likely other phonemes.

Strengths and Limitations

A notable strength of this study was that it attempted to control for several possibly confounding variables both between groups, within groups, and within the experimental stimuli. Between groups, this study controlled for age, gender, and most importantly, articulation ability. These controsl allowed conclusions to be drawn about how children with CIs acquire speech (i.e., covert contrast), and not simply describe the speech that they acquire. On the other hand, the variables that were controlled for within the experimental group (age of implantation and communication mode) and within the experimental stimuli (no place contrasts, stimuli shared coronal place of articulation) ensured that the chidren with CIs would be as successful as possible in the speech perception part of the OlimSpac

The first notable limitation of this study was the limited number of the stimulus items available to the listeners. The number of available stimulus items was limited by the number of times each phoneme was tested on the OlimSpac. Given the accuracy levels of all participants, there were many fewer /t/ errors than correct responses. In addition ,stiumulus selection was negatively impacted by the subjects' behavior. Several children from both groups were uncooperative during testing which affected vocal quality, intensity, intonation and background noise. This, in turn, reduced the overall number of speech samples with acceptable sound quality for listener rating. Excluding samples with poor quality made statistical analysis less robust, but did ensure that listeners judgements were based on the clearest exemplars of speech available and that their judgements weren't impacted by any non-speech noise.

The second major limitation of this study was that stimulus presentation was not randomized across blocks. Sounds in block 1, which presented the t/d contrast, were always rated first by listeners which may have resulted in different listener behavior between scales due to listener fatigue in the second block. This is partially supported by the data. Although there was no main effect observed for contrastive choice, visual inspection of the data revealed that in block 2, listeners were biased toward rating correct tokens of [t] from both groups as much less accurate than correct tokens of [t] in block 1.

The third limitation of this project is that it analyzed the characteristics of only three phonemes, and only one of the three in depth. Findings revealed interesting trends that may be indicative of idiosyncratic patterns of consonant acquisition. However, generalization of these findings to the development of speech in general in children with CIs is tenuous at best.

Clinical Implications

Children with CIs often receive intensive aural habilitation following implantation. During this therapy, these children would be best served by well-informed clinicians who are aware of their client's phonemic knowledge and gradual progress toward mastery of English consonants. As mentioned previously however, clinicians often record substituted and correct targets with identical phonetic symbols. This investigation has shown that, counter to customary clinical practices, distinctions among substitutions are often perceptible by an experienced listener under careful listening conditions. These findings have several implications for clinical practice involving prognosis, target selection, and therapy trajectory.

Prognosis. Children with CIs experience widely variable outcomes in aural habilitation. To date, research has been unable to identify a reliable characteristic that predicts successful CI use. However, covert contrast in NH children has been correlated with improved therapy outcomes. Since this investigation has tentatively established that children with CI covertly contrast certain consonants, it is possible that its presence may also predict outcomes in that population. A clinician who is able to reliably gauge the presence and extent of covert contrast may be able to provide more accurate prognostic statements.

Target selection. There are several schools of thought regarding target selection (Gierut, 2007; Miccio, 2005). Some believe selecting targets that are near mastery, or produced inconsistently, will provide better therapy results. Others favor target selections that the child has not yet acquired (i.e., the child is less stimulable). When choosing between these methods, a clinician must first assess the accuracy of target attempts. In this type of assessment, small phonetic differences are rarely considered. However, the listeners in this study were easily able to distinguish between correct and substituted target sounds. If a clinician wished to choose a sound that was close to mastery in order for the child to experience success, a substituted sound with covert contrast would be a better target. Further, if the clinician wished to target sounds for which the child had less phonological knowledge, she should choose sounds that the child produces with no contrast. Without considering the presence of covert contrast, a clinician might mistakenly assume that a child as no underlying form of a phoneme, when he does.

Therapy trajectory. Phonetic transcription is notoriously biased toward listener expectation. As a child progresses from producing no contrast, to a covert contrast, to an overt contrast, a clinician may be tempted to judge the phoneme as “mastered”, when it is still not adult-like. Using narrow transcription, Teoh and Chin (2009) reported that while a child may be perceived to produce an acceptable token of a particular consonant, he may still be perceived to have an “accent” (pg. 389), which can have profound psychosocial consequences if mainstream education is attempted. Conversely, a clinician may perceive that a child is failing to progress toward a contrast when progress is occurring, but has not crossed the threshold necessary for obvious perception. If gradient change in a child’s speech is not measured accurately, the target phoneme may be judged as accurate and therapy targets may be changed when not appropriate.

Directions for Future Research

Future studies should include more speakers, as well as a larger sample size of speech sounds to increase the robustness of statistical findings. Additionally, including several different phoneme contrasts, including a place-of-articulation contrast, could lend insight into how children with CIs develop speech in addition to simply describing the speech production of children with CIs. This project has shown that covert contrast is likely present in children with CIs as they gradually develop speech sounds, and that this phenomenon is perceptible to listeners. However, only one phoneme was examined, and for only two possible contrasts. Examining several other phonemes for covert contrast would greatly add to knowledge about how children with CIs acquire speech, and if they utilize the same processes as NH children.

Several recent studies have shown that early language learners (young children and older CI recipients) rely heavily on spectral information, such as formant transitions, to perceive phonetic contrasts (Nittrouer et al., 2009). It is well known that spectral information is poorly transmited by CI devices, regardless of speech processessing strategy. Therefore, it is not surprising that research has shown adult CI users who are sensitive to, and make perceptual use of spectral information generally, achieve higher word recognition scores than those who cannot (Moberly, et al., 2014). Sensitivity and attention to spectral information among CI users in the this study accounted for more variability than age of implantation. Investigating the pattern of covert contrast in children with CIs for other phoneme contrasts, and comparing it to what is typically seen in children with NH could shed light on how children with CIs (who have no NH language experience) utilize spectral cues.

These findings could be further supplemented by acoustic analysis of the speech sounds that were rated in this study. While it was clear that listeners perceived differences, it is impossible to describe the *nature* of the differences without some kind of instrumental aid. For example, [t] productions from the NH children in this study were found to be perceptually identical to t/□ substitutions, while the children with CIs produced them differently. With acoustic analysis, it may be possible to identify the acoustic-phonetic features that allowed listeners to differentiate the sounds of one group, but not the other. These findings might have implications for how children with CIs approach the acquisition task and whether or not it differs from NH children. Future reasearch should also explore the correlation between different acoustic cues and perceptions of accuracy in correct and substituted consonants. It is possible that significant differences exist in the speech of children with CIs and that they are more or less perceptible to listeners. This information could also inform therapy by allowing a clinician to selectively attend to features that are known to be disproportionately noticible to listeners.

References Cited

American Speech-Language-Hearing Association. (n.d.). *Birth to One Year*. Retrieved January 02, 2014, from What should my child be able to do?: http://www.asha.org/public/speech/development/01.htm

Avaaz Innovations, Inc. (2002). ECoS Experiment Generator 2.0 [computer software].

Blamey, P. J., Barry, J. G., & Pascale, J. (2001). Phonetic inventory development in young cochlear implant users 6 years postoperation. *Journal of Speech, Language, and Hearing Research, 44*, 73-79.

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer (version 5.3.56). The Netherlands: University of Amsterdam. Retrieved from www.praat.org October 2013.

Boothroyd, A., Eisenberg, L., & Martinez, A. (2006). *On-Line Imitative Test of Speech Pattern Contrast Production.* Los Angeles: House Ear Institute.

Caldwell, A., & Nittrouer, S. (2013). Speech perception in noise by children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *56*, 13-30.

Chin, S. B. (2003). Children's consonant inventories after extended cochlear implant use. *Journal of Speech, Language, and Hearing Research, 46*, 849-862.

Clark, G. (2003). *Cochlear Implants: Fundamentals and Applications.* New York: Springer.

Connor, C. M., Hieber, S., Arts, H. A., & Zwolan, T. A. (2000). Speech, vocabulary, and the education of chidlren using cochlear implants: Oral or total communication? *Journal of Speech, Language, and Hearing Research, 43*, 1185-1204.

Dillon, C. M., Cleary, M., Pisoni, D. B., & Carter, A. K. (2004). Imitation of nonwords by hearing-impaired children with cochlear implants: segmental analyses. *Clinical Linguistics and Phonetics, 18*, 39-55.

Dunn, D. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition.* Minneapolis: NCS Pearson, Inc.

Ertmer, D. J., & Goffman, L. (2011). Speech production accuracy and variability in young cochlear implant recipients: Comparisons with typically developing peers. *Journal of Speech, Language, and Hearing Research, 54*, 177-189.
Ertmer, D. J., True Kloiber, D., Jongmin, J., Connell Kirleis, K., & Bradford, D. (2012). Consonant production accuracy in young cochlear implant recipients: Developmental sound classes and word position effects. *American Journal of Speech-Language Pathology, 21*, 342-353.

Flipsen, P. (2011). Examining speech sound acquisition for children with cochlear implants using the GFTA-2. *The Volta Review, 11*, 25-37.

Gierut, J. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology, 16*, 6-17.

Giezen, M. R., Escudero, P., & Baker, A. (2010). Use of acoustic cues by children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 53*, 1440-1457.

Goldman, R., & Fristoe, M. (2000). *Goldman Fristoe Test of Articulation—Second Edition*. MN, Minneapolis: Pearson Assessments.

Gonzalez, V. (2013). *Effects of speech production ability on a measure of speech perception capacity in young children wiht cchlear implants and their articulation-matched peers, ProQuest Dissertaions and Theses* (Order No. 3604171, University of South Florida), 119.

Hicks, C. B., & Ohde, R. N. (2005). Developmental role of static, dynamic, and contextual cues in speech perception. *Journal of Speech, Language, and Hearing Research, 48*, 960-974.

Kreiman, J., Gerrat, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framewok for future research. *Journal of Speech and Hearing Reasearch, 36*, 21-40.

Lee, D. S., Lee, J. S., Oh, S. H., Kim, S. K., Kim, J. W., Chung, J. K., . . . Kim, C. S. (2001). Cross-modal plasticity and cochlear implants. *Nature, 409*, 149-150.

Li, F., Edwards, J., & Beckman, M. E. (2009). Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics, 37*, 111-124.

Macken M, & Barton D. (1980). The acquisition of voicing contrasts in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language,7*, 41–74.

Miccio, A. (2005). A treatment program for enhancing stimulability. In A. G. Kamhi, & K. E. Pollock (Eds.), *Phonological Disorders in Children* (pp. 163-173). Baltimore, MD: Paul Brookes Publishing Co.

Moberly, A. C., Lowenstein, J. H., Tarr, E., Caldwell-Tarr, A., Welling, D. B., Shahin, A. J., & Nittrouer, S. (2014). Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing? *Journal of Speech, Language & Hearing Research, 57*, 566-582.

Munson, B., Edwards, J., & Beckman, M. (2005). Phonological knowledge in typical and atypical speech–sound development. *Topics in Language Disorders, 25,* 190–206.

Munson, B., Edwards, J., Schellinger, S. K., Beckman, M. E., & Meyer, M. K. (2010). Deconstructing phonetic transcription: covert contrast, perceptual bias, and an extraterestrial view of Vox Humana. *Clinical Linguistics and Phonetics, 24*, 245-260..

Munson, B., Schellinger, S. K., & Carlson, K. U. (2012). Measuring speech-sound learning using visual analogue scaling. *SGI 1 Perspectives on Language Learning and Education, 19*(1), 19-30.

National Institute of Health and Human Services, National Institute of Health, National Institute on Deafness and Other Communication Disorders, NIDCD Office of Health Communication and Public Liaison. (2010). *Cochlear Implants Fact Sheet.* Retrieved from http://report.nih.gov/NIHfactsheets/Pdfs/CochlearImplants(NIDCD).pdf., June 2013.

Nittrouer, S., & Burton, L. T. (2001). The role of early language experience in the development of speech perception and language processing abilities in children with hearing loss. *The Volta Review, 103*, 5-37.

Nittrouer, S., Lowenstein, J. H., & Packer, R. R. (2009). Children discover the spectral skeletons in their native language before the amplitude envelopes. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1245-1253.

Pisoni, D. B., Cleary, M., Geers, A. E., & Tobey, E. A. (1999). Individual differences in effectiveness of cochlear implants in children who are prelingually deaf: New process measure of performance. *The Volta Review, 101*, 111-165.

Raviv, T., Kessenich, M., & Morrison, F. J. (2004). A mediational model of the association between socioeconomic status and three-year-old language abilities: The role of parenting factors. *Early Childhood Research Quarterly, 19*, 528-547.

Sarant, J. Z., Blamey, P. J., Dowell, R. C., Clark, G. M., & Gibson, W. R. (2001). Variation in speech perception scores among children with cochlear implants. *Ear & Hearing, 22,* 18-28.

Schellinger, S., Edwards, J., & Munson, B. (2010). *The role of intermediate productions and listener expectations onthe perception of children's speech.* Retrieved from: http://www.tc.umn.edu/~munso005/SchellingerEtAl_ToBeSubmitted_July8.pdf

Scobbie, J., Gibbon, F., Hardcastle, W. J., & Fletcher, P. (2000). Covert contrasts as a stage in the acquisition of phonetics and phonology. In M. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon* (pp. 194-207). Cambridge, U.K.: Cambridge University Press

Serry, T. A., & Blamey, P. J. (1999). A 4-year investigation into phonetic inventory development in young cochlear implant users. *Journal of Speech, Language and Hearing Research*, *42*, 141-154.

Sharma, A., Dorman, M. F., & Spahr, A. J. (2002). Rapid development of cortical auditory evoked potentials after early cochlear implantation. *NeuroReport, 13,* 1365-1368.

Shriberg, L. D., Gruber, F. A., & Kwiatkowski, J. (1994). Developmental phonological disorders III: Long-term speech normalization. *Journal of Speech and Hearing Research, 37*, 1151-1177.

Shu-Chen, P., Spencer, L., & Tomblin, J. (2004). Speech intelligibility of pediatric cochlear implant recipients with 7 years of device experience. *Journal of Speech, Language, and Hearing Research, 47*, 1227-1236.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*, 779-798.

SourceForge (2013). Audacity (Version 2.0.5)[Computer software].Retrieved from audacity. sourceforge.net/download October, 2013

Spencer, L. J., & Guo, L.-Y. (2013). Consonant development in pediatric cochlear implant users who were implanted before 30 months of age. *Journal of Deaf Studies and Education, 18*, 93-109.

Spencer, P. E. (2002). Language development of children with cochlear implants. In J. B. Christiansen, & I. W. Leigh, *Cochlear Implants in Children: Ethics and Choices* (p. 223-249). Washington D.C.: Gallaudet University Press.

Stoel-Gammon, C., & Otomo, K. (1986). Babbling development of hearing-impaired and normally hearing subjects. *Journal of Speech and Hearing Disorders, 51*, 33-41.

Teoh, A., & Chin, S. (2009). Transcribing the speech of children with cochlear implants: Clinical application of narrow phonetic transcription. *American Journal of Speech Language Pathology, 18*, 388-401.

Tomblin, J., Barker, B. A., Spencer, L. J., Xuyang, Z., & Bruce, J. G. (2005). The effect of age at cochlear implant initial stimulation on expressive language growth in infants and toddlers. *Journal of Speech, Language, and Hearing Research*, *48*, 853-867.

Tye-Murray, N., Spencer, L., & Woodworth, G. (1995). Aquisition of speech by children who have prolonged cochlear implant experience. *Journal of Speech and Hearing Research, 38*, 327-337.

Weismer, G. (1984). Acoustic analysis strategies for the refinement of phonological analysis. *ASHA Monographs, 22,* 30-52.

Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: Current designs and future possibilities. *Journal of Rehabilitation Research & Development, 45*, 695-730.

Yiu, E., & Ng, C. (2004). Equal appearing interval and visual analogue scaling of perceptual roughness and breathiness. *Clinical Linguistics & Phonetics, 18*, 211-229.

Appendix A

Demographic Characteristics of CI Participants

ID	Age (mo)	Gender	Race/ Ethnicity	Parent Education* Mat	Pat	Age at Implantation (mo)	Age at Activation (mo)	Experience with Device (mo)
CI01	70	F	H	3	4	21	22	48
CI02	65	M	W	4	4	8	9	55
CI03	56	F	W	4	2	14	15	40
CI04	43	F	B	2	2	24	26	16
CI05	42	M	W	3	4	18	19	22
CI06	76	F	W	2	1	21	21	55
CI07	70	M	W	2	1	18	20	50
CI08	35	M	H	3	4	7	8	26
CI09	59	M	B	1	999	29	30	28
Means	57					18	19	38

1 = High school diploma, 2 = Bachelor's Degree, 3 = Master's Degree/Graduate Certificate, 4 = Doctorate Degree, 999 = did not report

Appendix B

Matching Criteria for the Participants

	CI Participants						Articulation-Matched, Normal Hearing Participants				
Pairs	ID	Gender	Chron.-Age (mo)	Hearing-Age (mo)	GFTA-2 SS	GFTA-2 Confidence Interval (95%)	ID	Gender	Chron.-Age (mo)	GFTA-2 SS	GFTA-2 Confidence Interval (95%)
1	CI01	F	70	48	112	106-110	NH15	F	52	108	102-114
2	CI02	M	65	55	103	94-108	NH24	M	49	105	99-111
3	CI03	F	56	40	123	116-130	NH17	F	56	110	104-116
4	CI04	F	43	16	103	97-109	NH11	F	42	105	98-112
5	CI05	M	42	22	121	114-128	NH16	M	43	115	109-121
6	CI06	F	76	55	111	105-117	NH02	F	61	106	101-115
7	CI07	M	70	50	86	79-93	No match				
8	CI08	M	35	26	94	87-101	NH23	M	48	100	94-107
9	CI09	M	59	28	103	96-110	NH20	F	32	107	101-113

Appendix C

Subtest Scores from the OlimSpac for Each Participant

	CI Group												**NH Group**											
Match Pair	**AV Condition**						**AO Condition**						**AV Condition**						**AO Condition**					
	VH	**VP**	**CV**	**CC**	**CPf**	**CPr**	**VH**	**VP**	**CV**	**CC**	**CPf**	**CPr**	**VH**	**VP**	**CV**	**CC**	**CPf**	**CPr**	**VH**	**VP**	**CV**	**CC**	**CPf**	**CPr**
1	8	8	7	6	8	4	8	8	8	7	8	8	8	8	7	8	8	7	7	8	6	8	7	8
2	8	8	7	7	8	8	8	8	8	8	8	8	8	8	8	8	7	7	8	8	8	8	7	8
3	8	8	8	8	8	8	8	8	8	8	8	8	8	8	7	8	8	8	8	8	8	8	8	8
4	8	7	6	7	8	6	8	7	5	6	7	8	8	8	5	8	7	6	8	8	7	8	7	4
5	8	8	6	8	8	6	8	8	6	8	7	6	8	8	8	8	8	7	8	8	7	7	8	7
6	8	8	6	8	6	8	8	8	7	8	7	8	8	8	8	7	8	8	8	8	8	8	8	8
7	8	8	8	8	8	6	8	8	8	8	7	6	No Match											
8	8	7	6	7	6	8	8	8	8	7	5	7	8	8	7	8	8	4	8	8	8	8	7	6
9	8	8	5	4	7	6	8	6	6	4	7	7	8	8	8	7	8	7	8	8	6	8	8	6

Note. AO = Auditory-only, AV = Auditory-visual. Single Speech Contrasts: VH = vowel height, VP = vowel place, CV = consonant voicing, CC = consonant continuance, CPf = pre-alveolar consonant place, CPr = post-alveolar consonant place.

Appendix D

Listener Instructions for Practice and Experimental Blocks

Instructions for Practice Portion of Block 1:

You will hear a series of children producing VCV syllables that vary in quality. Listen for the consonant between the vowels. Your task is to identify whether a /t/ or /d/ phoneme was produced and to rate the quality of that phoneme production. Using the mouse, click the point on the scale that indicates how accurately the /t/ or /d/ was produced by the child. Marking the line near the ends means that you heard a "good" or nearly perfect /t/ or /d/. A mark in the middle means that the consonant sounds like a perfect mix between /t/ and /d/, you didn't recognize the sound, or you think it's a sound other than /t/ or /d/. Press OK to complete several practice trials.

Instructions for Experimental Portion of Block 1

The sounds you just head are very similar to the sounds that you are about to hear. Some were shorter or louder than others, some had background noise, the child's voice may rise or fall; ignore that. Focus only on each consonant. Remember to carefully rate the QUALITY of the production, and avoid simple identification. You may replay the sound as many times as you wish. Please consider the entire scale when ranking the quality of the production. At the end of part 1, you will have the opportunity to take a short break. Press OK to begin the experiment.

Instructions for Practice Portion of Block 2

For the next part of the experiment, you will again listen to children producing VCV syllables and rate the quality of the consonant produced. This time, however, you will choose between /t/ and /ch/. Again, a mark at either end of the scale will represent a perfect example of either a /t/ or /ch/. A mark in the middle will represent either a different phoneme, or a perfect blend between /t/ and /ch/.Press OK to complete several practice trials.

Instructions for Experimental Portion of Block 1

The sounds you just head are very similar to the sounds that you are about to hear. Carefully rate the QUALITY of the production, and avoid simple identification. You may replay the sound as many times as you wish. Please consider the entire scale when ranking the quality of the production. At the end of this part, the experiment will end. Press OK to begin.