COVERT CONTRAST IN THE EARLY DEVELOPMENT OF SPEECH SOUNDS IN CHILDREN USING COCHLEAR IMPLANTS

Ruth HUNTLEY BAHR[1], Terry GIER[2], Laura CONOVER[1]
[1]University of South Florida, Tampa, FL USA,
[2]Our Children's Academy of Winter Haven, Winter Haven, FL USA
rbahr@usf.edu, terry2@mail.usf.edu, lconover1@mail.usf.edu

Abstract
Covert contrasts represent intermediate productions that allow us broader insight into how children acquire a phonological system. However, little is known about the use of covert contrast in the development of speech sounds in children with cochlear implants (CIs). In particular, are these children using covert contrast in the same way that children with normal hearing (NH) do? Nine congenitally deafened children with CIs, ages 2;11 to 6;4 years (M = 4;9), who were implanted before age 3 were matched to typically developing children by articulation ability and gender. Their VCV productions from the OlimSpac were rated by 33 experienced listeners on an equal appearing interval scale to rate the phonetic accuracy of /t/ and its production as a substitution for /d/ and /ʧ/. Results indicated no differences in [t] production across groups. However, children with NH had a large, well-developed contrast between /t/ and /d/, but the later developing /tʃ/ showed little contrast with /t/. Children with CIs demonstrated the opposite trend. Their [t] for /d/ substitutions were more /t/-like, suggesting insufficient covert contrast for the voicing difference between these phones. However, they displayed a larger contrast for /t/ and /tʃ/ than the children with NH.

Keywords: speech production, speech sound development, cochlear implants, covert contrast, listener ratings

1 Introduction

Speech perception may be categorical, but speech production is not (Munson et al., 2010). Since human speech is characterized by inter- and intraspeaker variation, the precision of speech sounds can fall anywhere along the continuum between two phonemically similar sounds. This phenomenon rarely presents a problem in normal conversation, as listeners are biased to resolve ambiguous sounds to perceive words rather than nonwords (Strombergsson, Salvi, & House, 2015). However, a listener's strong sense of categorical perception begins to break down when he/she is asked to estimate the "goodness" of a sound production on a scale between two potential target phonemes, rather than using a simple forced-choice task (Strombergsson et al., 2015). While not always the

DOI: http://doi.org/10.18135/CAPSS.139

case, some of these ambiguous or less “good-fitting” sounds can be acoustically different from the target phonemes in ways that are detectible using spectrograms and other acoustic methods. As such, these ambiguous productions comprise a category of sounds known as covert contrasts, which are defined as “impressionistically homophonous [speech sound] categories that can be reliably distinguished at the phonetic level” (Kirby, 2011, p.1090).

In this paper, we will focus on the use of covert contrast in the development of speech sounds by children with cochlear implants (CIs) as one of several factors that influence speech development. Since each language has its own phonological contrasts, children must master specific perceptual and articulatory skills to become a proficient speaker of that language. In particular, infants need to learn which aspects of a sound function as unique cues for the production of meaningful speech. As young children develop phonetic categories, they go through a stage where their productions may be perceptually unreliable but acoustically distinct (i.e., covert contrasts). This type of production is often used as evidence against the traditional view that pronunciation shifts in children are caused by solely phonological changes (Scobbie et al., 1996; Strombergsson et al., 2015). As such, covert contrasts represent intermediate productions that are their own stage of learning, allowing researchers broader insight into how children acquire a phonological system (Hewlett & Waters 2004; Munson et al., 2012; Munson et al., 2017). Additional support for the use of covert contrast in speech development comes from the fact that children with speech sound disorders who produce covert contrasts have much better prognoses in treatment than those who do not (Byun et al., 2016).

1.1 Covert contrast

The fact that a child produces a covert contrast between two phonemes suggests that he/she can perceive some difference between them (Byun et al., 2016), which can then be refined into distinct phonemes. Research has shown that covert contrast has been observed for place of articulation for stops (Forrest et al., 1990), place of articulation for fricatives (Li et al., 2009), and voicing for stops (Macken & Barton, 1980). Of particular interest here is the research that describes factors that influence the perception of covert contrast. For instance, the context of the speech sound influences its perception. Sounds presented in real words are easier to detect than those presented in non-words (as reviewed by Strombergsson et al., 2015). These researchers also noted that the frequency of the phonotactic context and the listener’s level of experience also influences the perception of covert contrast. Finally, Munson et al. (2010) demonstrated that speaker age influences the perception of covert contrast. These researchers described how older speakers’ ambiguous phoneme productions were more likely to be judged as errors because the listener expected that elders should have well-developed phonemes.

The perception of subtle differences in speech sounds is essential for individuals who may receive a distorted or diminished speech signal, such as children who use cochlear implants (CIs). It has been well-documented that the signal delivered by CIs, although adequate for reasonably accurate speech perception, is significantly degraded in relation to the acoustic information that is available to a person with normal hearing. This is due to the processing methods common to CIs (Pisoni et al., 1999; Spencer, 2002). This modified speech signal might influence the speech features noted in the speech production of children who use CIs and might negatively impact the production of covert contrast.

1.2 Speech production in children with CI

There is significant variation in the speech production skills of children who use CIs. One of the predictors of speech accuracy is whether or not the child has successfully formed phonological representations of the speech sounds they are attempting to use (James et al., 2008). The production of covert contrast indicates that such phonological representations are developing, as the speaker is producing some acoustic difference between attempts that may be perceived as phonetically similar. Successful use of covert contrasts in children in the early stages of cochlear implant use would suggest they are likely to achieve better speech intelligibility than those who produce clear phoneme substitutions for longer periods of time.

Despite demonstrated variability in speech intelligibility, children with hearing loss show an initial accelerated growth in phoneme development after CI implantation, followed by a plateau where consonantal order of acquisition generally mirrors that of NH children, but at a slower rate (Blamey et al., 2001; Serry & Blamey, 1999; Spencer & Guo, 2013). This finding is more robust when device experience, as opposed to the chronological age of the child, is used as the metric for comparison with typically developing children (Flipsen, 2011). Nevertheless, some studies have suggested that the order of consonant acquisition in children with CIs differs slightly from that of typically developing children. Ertmer et al. (2012) found that some late-developing phonemes were produced more accurately than middle- or early-developing phonemes. Several other studies have identified that the /t/ productions of children with CIs were significantly less accurate than those of children with normal hearing (NH; Blamey et al., 2001, Chin, 2003; Ertmer et al., 2012). These same studies showed that production of /d/ was not similarly delayed. Additionally, the later-developing affricate /ʧ/ has been shown to emerge in children with CI significantly earlier than in children with NH (Ertmer et al., 2012; Spencer & Guo, 2013). Nevertheless, these trends have not been noted consistently, perhaps due to differences in research methodology, such as whether the researchers used broad or narrow transcription.

Given the noted differences in the speech production ability of children with CIs, it is likely that phonetic transcription alone will not adequately describe their early speech productions. This hypothesis led researchers to consider different measurement techniques. For instance, Schellinger and colleagues (2017) demonstrated that listeners could distinguish small, but statistically significant differences, in phonetic detail in children's speech when asked to rate productions on a visual analogue scale (VAS). Hence, increasing the depth of perceptual choice could produce a tool that can reliably reveal covert contrasts that listeners have been unable to identify using forced-choice or transcription measures alone. Such a finding is useful because assessing the presence of covert contrasts in speech productions holds clinical value (Munson et al., 2012). For instance, children who produce covert contrasts have a much higher likelihood of learning to correctly pronounce target phonemes than those who do not (Strombergsson et al., 2015), and children with speech-sound disorders who do not produce covert contrasts typically require longer treatment times (Tyler et al., 1993). Finally, the ability to measure the presence of covert contrast would imply the ability to track the progress of phoneme development from immaturity to maturity. Clinicians would not be forced into a choice of either correct or incorrect but would be able to track subtle changes during treatment.

1.3 Purpose of the study

As demonstrated, there is significant variation in speech production ability in children with CIs as they develop speech (Blamey et al., 2001; Ertmer & Goffman, 2011; Flipsen, 2011; Spencer & Guo, 2013). Previous research has identified and classified speech sound errors, created phonetic inventories to illustrate phonological knowledge, 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; Flipsen, 2011; Ertmer et al., 2012; Spencer & Guo, 2013). These studies found that, overall, children with CIs develop speech similarly to children with NH. However, some phonemes appear to develop in non-typical ways, and there is no clear explanation for this finding. Examination of covert contrasts in speech sounds produced by both children with CIs and children with NH can shed light on this issue and may have important clinical implications. 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 young children. Since /t/ has been repeatedly shown to be unusually late-developing in children with CIs compared to children with NH (Blamey et al., 2001; Ertmer et al., 2012; Spencer & Guo, 2013), it was chosen as the phoneme of interest in this investigation. With these factors in mind, there are two research questions that will be addressed:

> 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 [t] for another sound, are there significant perceptible differences (or covert contrasts) between the /t/ used as a substitution for /d/ or /tʃ/ and typical /t/ productions?

2 Methods

2.1 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 the 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.

2.1.1 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. 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 previous research has shown that children trained in oral communication have superior consonant acquisition when compared to children with CIs trained with other modes of communication (Connor et al., 2000). Table 1 lists the demographic characteristics of this group.

Table 1. Demographic Characteristics of CI Participants
*1 = High school diploma, 2 = Bachelor's Degree, 3 = Master's Degree/Graduate Certificate, 4 = Doctorate Degree, 999 = did not report
H = Hispanic, C = Caucasian, AA = African American.

ID	Age	Gender	Race/ Ethnic Group	Parent Education**		Age at Implantation (mo)	Age at Activation (mo)	Device Experience (mo)
				Mom	Dad			
CI01	70 mo.	F	H	3	4	21	22	48
CI02	65 mo.	M	C	4	4	8	9	55
CI03	56 mo.	F	C	4	2	14	15	40
CI04	43 mo.	F	AA	2	2	24	26	16
CI05	42 mo.	M	C	3	4	18	19	22
CI06	76 mo.	F	C	2	1	21	21	55
CI07	70 mo.	M	C	2	1	18	20	50
CI08	35 mo.	M	H	3	4	7	8	26
CI09	59 mo.	M	AA	1	999	29	30	28

2.1.2 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 child 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 (i.e., the sum of all articulation errors) were converted into a standard score 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 Table 2). 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." 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 achieved similar raw scores. Given these circumstances, they were considered to have similar articulation abilities and were paired.

Table 2. Matching Criteria for the Participants

	Participants with Cochlear Implants						Articulation-Matched, Normal Hearing Participants				
Pairs	ID	Gender	Chron. Age (mo)	Hearing Age (mo)	GFTA-2		ID	Gender	Chron. Age (mo)	GFTA-2	
					SS	95%CI				SS	95%CI
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	CI08	M	35	26	94	87-101	NH23	M	48	100	94-107
8	CI09	M	59	28	103	96-110	NH20	F	32	107	101-113

2.1.3 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.

2.2 Materials

Speech and language data were obtained using the Peabody Picture Vocabulary Test 4 (PPVT-4; Dunn & Dunn, 2007), the Goldman-Fristoe Test of Articulation-2 (GFTA-2; Goldman & Fristoe, 2000), and the On-line Imitative Test of Speech-Pattern Contrast Perception (OlimSpac; Boothroyd et al., 2010). The first two measures reflected speech and language ability. Speech samples were taken from the OlimSpac. This computerized software program provides a measure of speech perception by assessing the production of six phonologically significant speech contrasts in children with hearing loss (see Table 3).

During OlimSpac testing, pre-recorded VCV nonwords were presented over a loudspeaker, while the child was seated in front of a computer monitor in a sound-proof booth. The child was 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 to the child 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. Selected contrasts were consistent among subjects but presented in a random order during each test session. Each child imitated 16 VCV nonwords in each condition (auditory-visual, and auditory-only), for a total of 32 imitated productions per child. The children's imitated productions were recorded for future analysis using an Olympus ME52 directional lapel microphone connected to an RCA VR 5220 digital voice recorder. These productions served as the acoustic stimuli for the current investigation.

Table 3. 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. /uʧu/

2.2.1 Development of experimental protocol

For this project, a graduate student in speech-language pathology (SLP) phonetically transcribed subject responses, from the GFTA-2 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 and made the final decision. These transcriptions and OlimSpac recordings then were analyzed by the second author, who did not participate in the testing of the participants or scoring of the GFTA-2. She determined whether the VCV syllables represented a correct production or a clear substitution. Distortions were counted as correct, despite mild phonetic differences (inappropriate aspiration, imprecise production, etc).

The investigators selected /t/, /d/, and /ʧ/ as the phoneme productions of interest. These phoneme choices were particularly appropriate because one differed in voicing ([t] for /d/) and the other in manner of articulation ([t] for /ʧ/). 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] for /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, ʧ/ perception (Dillon et al., 2004).

For each subject, every opportunity for the three target consonants was isolated and digitized at 20,000 Hz using Praat (Boersma & Weenink, 2013). Each child had eight opportunities to produce /t/, and four opportunities each to produce both /d/ and /ʧ/. The following VCV contexts were utilized: /ata/, /utu/, /iti/, /ada/, /udu/, /iʧi/, and /uʧu/. 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.

The selected files underwent noise reduction using Audacity® (SourceForge, 2013) and were then normalized.

The prepared sound files were divided into two blocks: all samples of target [d] were placed into block 1, and all samples of target [ʧ] were placed into block 2. Samples of target [t] were equally distributed between the two blocks. Each block contained 60 unique speech production samples evenly distributed across CI and NH children. File order was 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.

Although previous studies on covert contrasts had used visual analogue scales (VAS), this experiment used equal-appearing interval scales (EAI scales). According to Yiu and Ng (2004), EAI scales showed significantly higher intra-rater reliability than VAS (EAI agreement = 0.73; VAS agreement = 0.57), and there was a moderate correlation (.56-.76) between EAI and VAS scale ratings for identical stimuli. Since consistent judgments are essential when assessing a child's progress toward a target sound, the use of an EAI scale should produce similar results to VAS and was used in this experiment.

2.3 Procedures

When the listeners arrived to participate in the study, they were asked to fill out a brief questionnaire 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.

ECoS Win experimental design software (Avaaz, 2002) was used to present the experimental trials on a Dell Optiplex desktop using Califone circumaural headphones. Each experimental block was preceded by a training block consisting of 10 novel sound files that were not utilized in the experimental blocks. The listeners were told that they would be listening to children producing VCV nonwords and were given an example (like [ada]). Then, listeners were shown a 7-point EAI scale. They were asked 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 a 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. In block 1, listeners rated the subjects' attempts at /t/ and /d/. In block 2, they rated attempts at producing /t/ and /ʧ/.

3 Results

3.1 Intra-rater reliability

Over both experimental blocks, each listener rated 12 stimuli twice (N = 48 trials). For each listener, the percentage of responses to duplicated stimuli that were within ± 1 scale value of the original rating was calculated (Kreiman et al., 1993). These values were averaged across listeners. Calculations revealed that overall, 88.1% of duplicated trials were within ± 1 scale value of the original rating. Of these, 56.6% were in exact agreement. Hence, listener reliability was determined to be very good.

3.2 Statistical analyses

A three-way repeated measures ANOVA was conducted to analyze the influence of group (CI vs NH), Transcription Category (4 levels of correct and substituted productions), and Covert Contrast Category (/d/ or /ʧ/) on perceptual ratings. Results revealed a significant three-way interaction. However, differences across the experimental blocks were not of primary interest and will not be discussed further. Statistical analysis also 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 found across both groups and TC. However, the statistically significant interaction between group and TC was of particular importance, therefore, the research questions will be addressed within this interaction.

3.2.1 Accuracy of /t/ productions

In the past, /t/ has been shown to be unusually late developing in children with CIs (Blamey et al., 2001; Chin, 2003; Ertmer et al., 2012). The first goal of this project was to confirm this observation by examining listener perceptions of [t] 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 out of 8 paired comparisons of interest were not significant: [t] for /t/ in both experimental blocks, and [ʧ] for /ʧ/ (see Figures 1 and 2). In other words, [t] and [ʧ] productions were similarly accurate across groups; however [d] for /d/ productions were significantly more accurate in children with NH. Hence, when the production was judged to be a /t/ by SLPs, children with CIs successfully produced /t/ as accurately as their NH peers.

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 [t] productions from NH children. To test this hypothesis, overall error frequency taken from the OlimSpac testing was determined to provide enough additional relevant information to warrant analysis. A confusion matrix of CI group productions had previously been generated when selecting contrastive consonant choices. To compare error frequencies between groups, a second confusion matrix (of NH productions) was created (see Table 4).

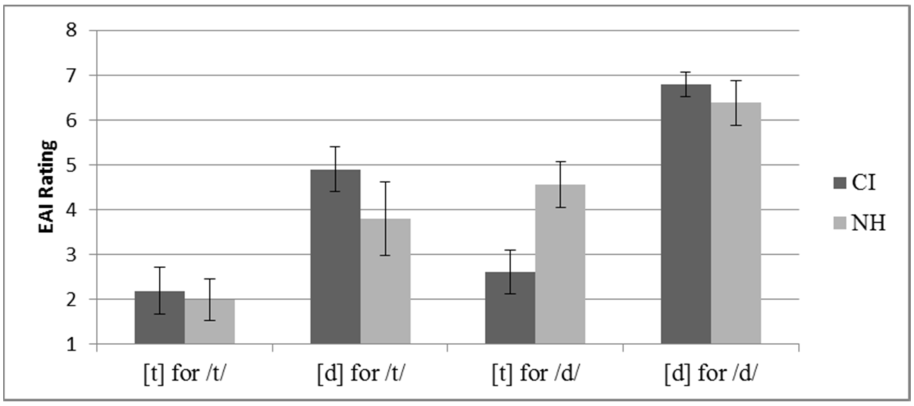

Figure 1.
Differences in listener perceptions of consonant accuracy
for [t] for /t/ and [d] for /d/.
*Covert contrast is shown in the [d] for /t/ and [t] for /d/ contrasts.

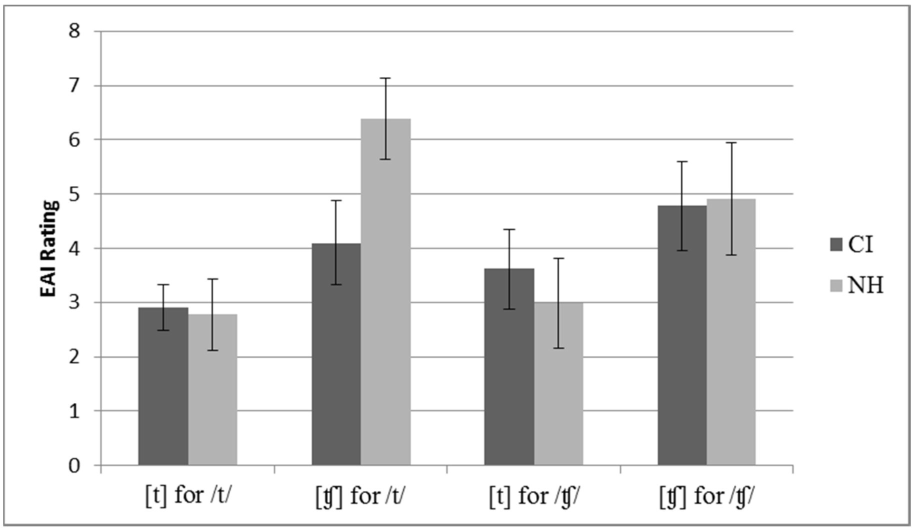

Figure 2.
Differences in listener perceptions of consonant accuracy
for [t] for /t/ and [ʧ] for /ʧ/.
*Covert contrast is shown in the [ʧ] for /t/ and [t] for /ʧ/ contrasts.

Table 4. Confusion matrix for responses on the OlimSpac produced by children with CI vs. NH

	Child's Production **CI(top) NH(bottom)**							
OlimSpac Target	/p/	/b/	/d/	/t/	/s/	/ʃ/	/ʧ/	**Other**
/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 12.50				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
/ʃ/			5.56	5.56	2.78 21.88	66.67 71.88	13.89 3.13	5.56 3.13
/ʧ/			2.78	19.44 21.88		22.22 6.25	55.56 71.88	

Examination of this confusion matrix revealed that children with NH produced /t/ accurately in 87.93% of opportunities, whereas children with CIs produced it accurately in 73.24% of opportunities. Hence, [t] was found to be perceptually less accurate in children with CI when compared to those [t]s produced by children with NH. Nevertheless, [t] was the most accurate phoneme produced by the children with CI when compared to the other OlimSpac test stimuli, which also matched the performance of the NH group.

3.2.2 Perceptible contrasts 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, were the patterns of covert contrast similar to those observed in children with NH? Identification of covert contrast was best addressed by an 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 rated significantly different from the [t] used as a substitution, as well as contrasting with [d] and [ʧ] when they were used as a substitute for a /t/. In addition, post hoc testing revealed significant differences in the similarity of [t, d, ʧ] productions across groups when they were used as substitutions for other phonemes (i.e., [t] for /d/, [d] for /t/, [t] for /ʧ/, [ʧ] for /t/). This finding suggests that there were significant differences in the patterns of covert contrast across groups. As illustrated in Figures 1 and 2, all four paired comparisons involving phoneme substitutions across groups were significant ($p < .001$). When children with CIs substituted [d] for /t/, it was perceived as more [d]-like and when they substituted [t] for /d/, it was perceived as more [t]-like. The opposite pattern was noted in NH children. However, a different tendency was noted for /ʧ/. For children with CIs, the [t] for /ʧ/ substitution was more /ʧ/-like than for NH hearing children. While there was a significant group difference for the [ʧ] for /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 in this condition for the children with CIs was more [ʧ]-like.

4 Discussion

The current results suggest the measurement technique used by the listener does influence the reporting of developmental speech patterns for children with CIs. When using phonetic transcription, the children with CIs were less accurate in phoneme production than speech age-matched children with NH. However, when EAI scales were used to rate the same speech productions, listeners identified different patterns of covert contrast across these groups.

4.1 The development of /t/ in children with CIs

The first research question dealt with the accuracy of /t/ production when the listener decision of phonetic accuracy of /t/ across speaker groups (CI vs. NH) varied by technique: EAI scale versus phonetic transcription. A three-way repeated measures ANOVA using the data from the EAI scale revealed that listeners perceived no significant difference between groups when only the correct /t/ productions were considered. Hence, children using CIs were no less accurate in their [t] productions than children with NH when speakers were matched for articulation ability.

These non-significant findings are likely related to advances in CI speech processing technology, as the previous studies that showed delayed /t/ acquisition were conducted over 10 years ago (Blamey et al., 2001; Chin, 2003). Another possible explanation involves device experience. The two previous studies that revealed delayed /t/ development tested participants with less than 3 years of device experience (Ertmer et al., 2012; Spencer & Guo, 2013). The children who participated in this project averaged three years of device experience. Given that children with CIs acquire speech sound accuracy quickly at first, and then slow down, it is possible that our participants were in the “plateau” stage, given their length of device experience and history of oral language use, whereas those in the comparison studies were still in the early stages of development, characterized by rapid growth in their phonetic inventories.

The second analysis of the accuracy of /t/ was derived from an examination of all phonemes tested on the OlimSpac. Error proportions for each participant were collapsed by group and placed in a confusion matrix. Results indicated that children with CIs made more speech sound errors than children with NH for all phones tested, including /t/. However, /t/ was not significantly more impaired than the other phonemes produced by children with CI. Interestingly, both groups produced /t/ accurately more often than other phonemes evaluated on the OlimSpac (e.g., /p, b, d, s, ʃ, ʧ/). These findings do not support those of Ertmer et al. (2012) who reported that initial /t/ was less accurate than both /d/ and /ʧ/ in children with CIs during acquisition.

4.2 Use of covert contrast

The second research question addressed the presence of covert contrast in children with CIs. To address this issue, listener ratings of /t, d, ʧ/ substitutions were compared with ratings of correct tokens for the same phonemes. Covert contrast was present if the two sounds (one substituted, one correct) were transcribed identically but rated differently by listeners on the EAI scale. If covert contrast was present, then the child was in the process of developing the speech sound. If not, then the error suggested lack of phonological knowledge for the target phoneme contrast.

Post-hoc comparisons of the group by transcription category (TC) interaction showed that both CI and NH groups produced perceptible differences between correct productions and substitutions for the target phonemes (Figures 1 and 2). The voicing contrast was more readily perceived in the productions by children with NH while the children with CIs struggled with this contrast. That is, the [t] for /d/ substitutions produced by children with CIs sounded more like [t] and [d] for /t/ substitutions sounded like [d]. These errors lack covert contrast and support difficulties with voicing. This finding confirms Gonzalez's (2013) conclusion that the children with CIs struggled with the perception and/or production of voicing more than with other phoneme distinctions. Since the OlimSpac uses VCV syllables, one might expect more difficulties with syllables that alternate voicing (/ata/) than one that is entirely voiced (/ada/). It was surprising that children with CIs struggled with both syllable types.

A comparison could not be made across groups for the manner (plosive/affricate) contrast since so few [ʧ] for /t/ errors were noted in the NH group. However, the children in the CI group produced a sufficient number of both [t] for /ʧ/ and [ʧ] for /t/ errors for analysis. Results indicated that those with CIs had good phonological representations for /t/, and the production of covert contrast in the errors revealed productions closer to the desired target, either [t] or [ʧ]. This demonstration of covert contrast in children with CIs supports the idea that they have acquired both [t] and [ʧ] but have not completely mastered either.

An interesting finding was that the children with CIs were able to produce a perceptible contrast between correct [t] and the [t] for /ʧ/ substitutions, while children with NH did not. As expected for children with NH who were approximately 4;0 years old, they did not have a mature phonological representtation for /ʧ/, as it is a later developing phoneme. On the other hand, unlike their NH peers, the children in the CI group, with an average chronological age of 4;9 years, were developing this contrast. Even though the CI group only had an average of 3;2 years of robust hearing experience, they were at the approximate chronological age for the development of /ʧ/ (Smit et al., 1990). Since these results are based on listener perceptions of covert contrast, it is possible that children with CIs use a certain speech feature (like aspiration or voice onset time) to make [t] substitutions sound more like [t] when contrasted with /d/, and more like [ʧ] when attempting to produce /ʧ/. Hence, the use of a CI might influence which acoustic cues the child attends to in the development of phonemic contrasts, or it is possible that these children weigh the available cues in a different way than children with NH do. More detailed acoustic analyses are needed to test these hypotheses.

4.3 Clinical implications

This investigation has shown that subtle differences in phoneme accuracy are often perceptible by an experienced listener. A clinician who is able to reliably gauge the presence and extent of covert contrast may be able to provide more accurate prognostic statements and select treatment targets that will facilitate student progress.

There are two different ways to select a target for children with NH (Gierut, 2007; Miccio, 2005). Based on the child's learning style, the clinician can choose a target for which the child has contrastive knowledge (i.e., a sound produced with covert contrast) or one that is unknown to the child. In other words, is phonetic accuracy or the learning of a new phonemic contrast the focus of treatment? Since research has demonstrated the utility of narrow transcription in the identification of speech sound errors in children with CIs (Teoh & Chin, 2009), it may be possible to incorporate an assessment of covert contrast in an evaluation of speech sound disorder so that treatment decisions can be enhanced. The current study indicates that covert contrast can provide the data necessary to make decisions about target selection and that covert contrast can be used to track progression towards phoneme mastery.

Acknowledgments

This paper is based on the Master's thesis of the second author.

References

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. S., & Martinez, A. S. (2010). An on-line imitative test of speech-pattern contrast perception OlimSpac: Developmental effects in normally hearing children. *Journal of Speech, Language, and Hearing Research, 53,* 531-542.

Byun, T. M., Buchwald, A., & Mizoguchi, A. (2016). Covert contrast in velar fronting: An acoustic and ultrasound study. *Clinical Linguistics & Phonetics, 30,* 249-276.

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

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.

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. A. (2011). Speech production accuracy and variability in young cochlear implant recipients: Comparisons with typically developing age-peers. *Journal of Speech, Language, and Hearing Research, 54,* 177-189.

Ertmer, D. J., Kloiber, D. T., Jung, J., Kirleis, K. C., & 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.

Forrest, K., Weismer, G., Hodge, M., Dinnsen, D. A., & Elbert, M. (1990). Statistical analysis of word-initial /k/ and /t/ produced by normal and phonologically disordered children. *Clinical Linguistics & Phonetics, 4,* 327-340.

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.* Minneapolis, MN: Pearson Assessments.

Gonzalez, V. (2013). *Effects of Speech Production Ability on a Measure of Speech Perception Capacity in Young Children with Cochlear Implants and their Articulation-Matched Peers.* Unpublished dissertation completed at the University of South Florida.

Hewlett, N., & Waters, D. (2004). Gradient change in the acquisition of phonology. *Clinical Linguistics and Phonetics, 18,* 523-533.

James, D., Brinton, J., Rajput, K., & Goswami, U. (2008). Phonological awareness, vocabulary, and word reading in children who use cochlear implants: Does age of implantation explain individual variability in performance outcomes and growth? *Journal of Deaf Studies and Deaf Education, 13,* 117-137.

Kirby, J. (2011). Modeling the acquisition of covert contrast. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1090-1093). Hong Kong.

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

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

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.

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

Munson, B., Schellinger, S. K., & Carlson, K. U. (2012). Measuring speech-sound learning using visual analog scaling. *SIG 1 Perspectives on Language Learning and Education, 191,* 19-30.

Munson, B., Schellinger, S. K., & Edwards, J. (2017). Bias in the perception of phonetic detail in children's speech: A comparison of categorical and continuous rating scales. *Clinical Linguistics & Phonetics, 31,* 56-79.

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 measures of performance. *The Volta Review, 101,* 111-164.

Schellinger, S. K., Munson, B., & Edwards, J. (2017). Gradient perception of children's productions of /s/ and /θ/: A comparative study of rating methods. *Clinical Linguistics & Phonetics, 31,* 80-103.

Scobbie, J. M., Gibbon, F., Hardcastle, W. J., & Fletcher, P. (1996). Covert contrast as a stage in the acquisition of phonetics and phonology. *QMC Working Papers in Speech and Language Sciences, 1,* 43-62.

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.

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, P. E. (2002). Language development of children with cochlear implants. In P. E. Spencer (Ed.), *Cochlear Implants in Children: Ethics and Choices* (pp. 222-249). Washington, D. C.: Gallaudet University Press.

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 Deaf Education, 18,* 93-109.

Strombergsson, S., Salvi, G., & House, D. (2015). Acoustic and perceptual evaluation of category goodness of /t/ and /k/ in typical and misarticulated children’s speech. *Journal of the Acoustical Society of America, 137,* 3422-3435.

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.

Tyler, A. A., Figurski, G. R., & Langsdale, T. (1993). Relationships between acoustically determined knowledge of stop place and voicing contrasts and phonological treatment progress. *Journal of Speech and Hearing Research, 36,* 146-759.

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