Interdependence of Linguistic and Indexical Speech Perception Skills in School-Aged Children with Early Cochlear Implantation

Ann Geers, Ph.D,
The University of Texas at Dallas

Lisa Davidson, Ph.D.,
Washington University in St. Louis

Rosalie Uchanski, Ph.D., and
Washington University in St. Louis

Johanna Nicholas, Ph.D.
Washington University in St. Louis

Abstract

Objectives—This study documented the ability of experienced pediatric cochlear implant (CI) users to perceive linguistic properties (what is said) and indexical attributes (emotional intent and talker identity) of speech, and examined the extent to which linguistic (LSP) and indexical (ISP) perception skills are related. Pre-implant aided hearing, age at implantation, speech processor technology, CI-aided thresholds, sequential bilateral cochlear implantation, and academic integration with hearing age-mates were examined for their possible relationships to both LSP and ISP skills.

Design—Sixty 9–12 year olds, first implanted at an early age (12–38 months), participated in a comprehensive test battery that included the following LSP skills: 1) recognition of monosyllabic words at loud and soft levels, 2) repetition of phonemes and suprasegmental features from non-words, and 3) recognition of keywords from sentences presented within a noise background, and the following ISP skills: 1) discrimination of male from female and female from female talkers and 2) identification and discrimination of emotional content from spoken sentences. A group of 30 age-matched children without hearing loss completed the non-word repetition, and talker- and emotion-perception tasks for comparison.

Results—Word recognition scores decreased with signal level from a mean of 77% correct at 70 dB SPL to 52% at 50 dB SPL. On average, CI users recognized 50% of keywords presented in sentences that were 9.8 dB above background noise. Phonetic properties were repeated from non-word stimuli at about the same level of accuracy as suprasegmental attributes (70% and 75%, respectively). The majority of CI users identified emotional content and differentiated talkers significantly above chance levels. Scores on LSP and ISP measures were combined into separate principal component scores and these components were highly correlated (r = .76). Both LSP and ISP component scores were higher for children who received a CI at the youngest ages, upgraded to more recent CI technology and had lower CI-aided thresholds. Higher scores, for both LSP and ISP components, were also associated with higher language levels and mainstreaming at younger ages. Higher ISP scores were associated with better social skills.

Corresponding Author: Ann Geers, Callier Center for Advanced Hearing Research, University of Texas at Dallas, 1966 Inwood Rd., Dallas, TX, 75235. Phone: 828-627-1697. Fax: 214-905-3146. ageers@utdallas.edu.

Conclusions—Results strongly support a link between indexical and linguistic properties in perceptual analysis of speech. These two channels of information appear to be processed together in parallel by the auditory system and are inseparable in perception. Better speech performance, for both linguistic and indexical perception, is associated with younger age at implantation and use of more recent speech processor technology. Children with better speech perception demonstrated better spoken language, earlier academic mainstreaming, and placement in more typically-sized classrooms (i.e., >20 students). Well-developed social skills were more highly associated with the ability to discriminate the nuances of talker identity and emotion than with the ability to recognize words and sentences through listening. The extent to which early cochlear implantation enabled these early-implanted children to make use of both linguistic and indexical properties of speech influenced not only their development of spoken language, but also their ability to function successfully in a hearing world.

Keywords

Cochlear Implant; pediatric hearing loss; speech perception; indexical

---

Perception of spoken language is facilitated by two complementary types of information in the speech signal: linguistic and indexical. Linguistic information conveys the meaning of the message or "what is said" while indexical information contributes information from or about a speaker or "how it is said" and "who said it." Linguistic and indexical attributes of the speech signal are simultaneously encoded in the speech waveform and are highly interdependent in normal-hearing listeners (Pisoni 1997). Linguistic attributes are encoded in discrete units such as phonemes or syllables which are necessary requisites for language reception. Indexical attributes include characteristics of the speaker such as age, gender, and dialect, as well as changing states of the speaker such as emotional mood or fatigue. Both types of information are part of the mental representation of spoken utterances and are crucial for conveying meaning in social spoken communication (Abercrombie 1967).

Recent studies suggest that the processing of linguistic and indexical information is interwoven in a complex way in individuals with normal hearing (Johnson et al. 2011; Perrachione et al. 2011). Rather than "normalizing" across talker variations to abstract common phonetic properties, evidence suggests that the listener integrates linguistic and indexical properties of the speech signal. Pisoni (1997) argues for a conceptual link between symbolic linguistic properties of speech and the simultaneously-encoded vocal source. In this view, the listener uses indexical properties of the vocal source such as gender, emotion and speaking rate to facilitate a phonetic interpretation of the linguistic content of the message. The extent to which cochlear implants (CIs) enable children to make use of both linguistic and indexical properties of speech may influence not only their development of spoken language, but also their ability to function successfully in a hearing world.

Measuring Linguistic Speech Perception in Children with CIs

Word Recognition

The most common method of assessing speech perception employs word lists presented in an open-set as pre-recorded or live-voice stimuli. The examiner assesses whether the child's spoken repetition is an acceptable rendition of the target stimulus. Recognition of spoken words is influenced not only by how well a child's device conveys the speech signal, but also by child characteristics, including linguistic and phonological processing skills. A longitudinal study of 112 CI users reported average word recognition scores improved from 50% to 60% correct on the Lexical Neighborhood Test (LNT; Kirk et al. 1995) between 5 and 13 years of CI experience (Davidson et al. 2011). This change in word recognition with longer CI experience reflects improved lexical access, speech articulation and use of

linguistic context (Davidson et al. 2011; Blamey, Sarant, Paatsch et al. 2001) as well as improved cognitive skills (Pisoni et al. 2011), all of which affect the child's ability to accurately recognize and respond to word lists (Eisenberg et al. 2002). A performance advantage has been documented for lexically "easy" words (high frequency-of-occurrence words from sparse lexical neighborhoods) compared to lexically "hard" words (low frequency-of-occurrence words from dense neighborhoods), and this advantage has been shown to increase with age (Davidson et al, 2011), suggesting that deaf children with CIs, like NH children, recognize spoken words relationally in the context of other spoken words they know in their mental lexicons.

Non-word Repetition

Use of nonsense words as stimuli greatly reduces the influence of lexical knowledge on speech perception scores. Like responses to 'real' words, non-word repetitions reflect a child's ability to hear speech sounds, but the child cannot rely on preexisting phonological representations as in a word-recognition test (Cleary et al. 2002). Correct imitation of a non-word, such as *ballop,* requires not only reproducing the individual phonemes (segmental features), but also the correct number of syllables (i.e., two) and appropriate stress (i.e., first syllable), both of which are supra-segmental features. The child's response reflects the ability to construct a new phonological representation based on their perception of both segmental and supra-segmental attributes of the novel auditory stimulus. This representation must then be maintained in short-term memory using verbal rehearsal until it is translated into an articulatory program for production. Therefore, in addition to assessing the perception of speech sounds, non-word repetition responses also reflect phonological processing, short-term memory, articulatory planning and speech production skill (Dillon et al. 2004). The task has been used to study phonological processing and speech production skills of children with language-learning difficulties (Botting & Conti-Ramsden 2001) as well as children with cochlear implants (Carter et al. 2002; Dillon et al. 2004). Previous research examining non-word repetition responses in CI users has demonstrated consonant imitation far worse than that of normal-hearing age-mates (Dillon, Pisoni, Cleary et al. 2004; Dillon, Cleary, Pisoni et al. 2004). Furthermore, non-word repetition has been used to measure CI children's facility with suprasegmental properties of speech (imitating syllable number and stress), revealing scores closer to those of normally-hearing (NH) children than their scores on segmental attributes (Carter et al. 2002). Use of non-word repetition as a speech perception measure is a relatively novel application of this technique, which has shown that non-word repetition scores (both segmental and suprasegmental) show strong associations with scores on standard word recognition tests (Cleary et al. 2002; Carter et al. 2002).

Measuring Speech Perception in Challenging Listening Conditions

Successful educational integration of children with CIs into regular classrooms with NH age-mates (i.e., mainstreaming) may be associated with strong linguistic speech perception skills, particularly in challenging listening conditions. Most everyday listening occurs in settings where audibility and speech recognition are degraded. For example, listening to speech with a +10 dB signal-to-noise ratio (SNR), compared to a quiet background, decreased average word recognition scores by 28 percentage points (81% vs. 53%) for a group of 112 adolescent CI recipients (Davidson et al. 2011). In addition, pediatric CI recipients require a higher SNR to achieve 50%-correct speech understanding (SNR-50) than the level documented by Killion et al. (2004) for NH children. The average SNR-50 on the Bamford-Kowal Bench-Speech In Noise test (BKB-SIN; Killion et al. 2004) was 10 dB for thirty pediatric CI recipients age 7–17 years (Davidson et al. 2009), about 10 dB higher than the SNR-50 for similarly-aged NH children. Perceiving speech at soft levels is also a challenge for both adults and children using CIs (Skinner et al. 1994; Dawson et al. 2004;

Firszt et al. 2004;). Listening to speech at a soft level (50 dB SPL), compared to a loud conversational level (70 dB SPL), decreased average word recognition scores by 20 percentage points in a study of 26 children (7–15 years old) with CIs (Davidson 2006) and by 13 percentage points in a different study of 112 teenage children (15–18 years old) using CIs (Davidson et al. 2011). Such performance decrements are not seen in children without hearing loss, who reach ceiling levels in word recognition at about 45 dBA (Eisenberg et al. 2002).

Measuring Indexical Perception in CI Users

The enormous variability in speech perception scores of children with CIs, even after several years of CI experience, suggests the need for greater understanding of underlying explanatory mechanisms. Cochlear implants provide deaf individuals with only limited access to the acoustic characteristics associated with indexical properties of the speech signal such as pitch or information about the temporal fine structure (TFS) of the speech waveform. While the benefits of CIs for perception of linguistic content have been well-documented (Cheng et al. 1999; Bond et al. 2009) using scores from standard speech tests, studies of indexical perception are fewer, though growing, in number (Spahr & Dorman 2004; Fu et al. 2004; Fu et al. 2005; Luo et al. 2007; Cullington & Zeng 2011). Cochlear implant processers have been designed to maximize segmental and linguistic perception of spoken language, particularly English and other non-tonal languages. Indexical information may not be as well represented. Indexical features are generally thought to be transmitted via the speech signal's time-intensity envelope and fundamental frequency information (Banse & Scherer 1996; Grant & Walden 1996; Bachorowski & Owren 1999; Scherer 2003). Listeners using a CI alone may have limited ability to perceive indexical features due to the reduced spectral resolution of current CI systems and reduced access to periodicity cues (Fu et al. 2004; Fu et al. 2005; Schvartz & Chatterjee 2012).

Perception of Emotional Content in Speech

CI users, whether children, adolescents or adults, generally have difficulty identifying vocal expressions of emotion (Luo et al. 2007; Hopyan-Misakyan et al. 2009; Most & Aviner 2009; Nakata et al. 2012). Hopyan-Misakyan et al. (2009) examined emotion recognition for 18 pediatric right-ear CI users (age 7–13 years) and 18 age- and gender-matched NH children. Recognition of emotion (happy, sad, angry, or fearful) in speech and in facial affect was examined using the Diagnostic Analysis of Nonverbal Behavior tasks (Nowicki & Duke 1994). Children with CIs performed more poorly than NH children on identifying emotion in speech (50% vs. 79% correct, respectively), but performed similarly on identifying emotion in faces (87% vs. 85% correct, respectively). Overall performance was not associated with age at test or time since CI activation. The authors concluded that while children with CIs are able to recognize emotion in faces their ability to recognize emotions in speech is limited.

Talker Discrimination

Cochlear implant users also have difficulty differentiating talkers (Fu et al. 2004; Cleary et al. 2005; Fu et al. 2005; Kovačić & Balaban 2009, 2010). In one of the earliest reports, Cleary and Pisoni (2002) evaluated the abilities of two groups of children to discriminate female talkers. Eight- and nine-year-old prelingually deaf children with CIs and 5-year-old NH children were presented pairs of sentences spoken by three females. Participants were asked to determine whether the sentences were spoken by the "same person" or by "different people." Two conditions were examined: i) 'fixed' in which the linguistic content of the two sentences (sentence script) was the same, and ii) 'varied' in which the linguistic content of the two sentences differed. Mean scores for children with CIs were higher in the 'fixed' than

in the 'varied' condition, and were substantially lower than scores for the younger NH children. In fact, in the 'varied' condition, most (> 80%) of the CI participants did not perform significantly different from chance. However there was considerable variability among children and talker discrimination performance was significantly correlated with word recognition scores, suggesting that acoustic cues for linguistic processing were also useful for indexical perception.

Cleary, Pisoni & Kirk (2005) compared perception of voice similarity in 5–12 year old CI users and NH 5-year olds to determine the minimum difference in fundamental frequency needed to distinguish different talkers. While NH children exhibited well-defined perceptual boundaries for discriminating within-talker from between-talker utterances, most of the experienced CI users were unable to do so. Variability within the CI group was large, however, and those who scored above 64% correct in open-set word recognition scored more like the children with NH on the talker discrimination task.

Two studies by Kovačič and Balaban (2009, 2010) examined the ability of 41 pediatric (5–18 yrs old) CI users to both identify and discriminate talker gender. More than half of the children (23 of 41) could not identify gender. The remaining children performed better than chance, but still identified gender much more poorly than even younger NH children (84% vs. 98% correct). However, some CI users could discriminate but not identify gender, indicating that acoustic cues may have been available through their implants, but long-term categorical memory of voice cues for gender may not be developed in these children. In the second study (Kovačič & Balaban 2010), gender identification performance was found to be significantly and inversely related to duration of deafness before cochlear implantation and to age at implantation. Comparing an individual voice with a stored model of male and female voices requires both short- and long-term memory resources, which may be disrupted by a long period of deafness or delayed onset of auditory input. Children who were deaf during a critical period for encoding auditory attributes such as pitch and timbre may be unable to relate perceived acoustic differences to long-term categorical representations.

Implications for Further Research

These recent results are consistent with a long-standing body of research indicating that indexical variations affect memory for spoken words, and that storage and processing of indexical information may have consequences for processing of the linguistic attributes of speech as well (e.g., Craik & Kirsner 1974; Goldinger et al. 1991; Pisoni 1997; Bradlow et al. 1999; Sidtis & Kreiman 2012). Collectively, these studies suggest that for listeners with normal hearing, the acoustic properties of a talker's voice, such as gender, dialect or speaking rate, are perceived and encoded in memory along with the linguistic message and are subsequently used to facilitate a phonetic interpretation of the linguistic content of the message. This process may be disrupted in CI users. While various developments in CI technology have improved linguistic speech perception abilities for CI recipients, indexical perception may not benefit. There are many documented increases in word recognition scores ascribed to changes in speech coding strategies, e.g., from M-Peak to SPEAK for Nucleus devices (Skinner et al. 1994; Skinner et al. 1996, 1999; Geers et al. 1999) and from compressed analogue to CIS for Advanced Bionics devices (Wilson et al. 1991). More recently, developments in automatic gain control, preprocessing strategies, and increased input dynamic range of speech processors have improved aided thresholds and perception of soft speech for adults and children (Cosendai & Pelizzone 2001; James et al. 2002; McDermott et al. 2002; Zeng et al. 2002; Donaldson & Allen 2003; James et al. 2003; Dawson et al. 2007; Holden et al. 2007; Santarelli et al. 2009; Davidson et al. 2009; Davidson et al. 2010). Yet, there are no known reports of the effect of improvements in CI device technology on the perception of indexical attributes of speech.

If indexical perception is limited by relatively poor low-frequency spectral resolution in CI processors then linguistic and indexical perception might be presumed to develop unevenly in CI users. In this case, the ability to recognize phonemes, words and sentences would be highly interrelated, but perhaps unrelated to the ability to discriminate vocal characteristics such as gender and emotion. On the other hand, if indexical perception is related to cognitive processing mechanisms that also underlie perception of linguistic cues, then these skills might be expected to develop in tandem, and share common predictors. For example, like voice discrimination (Kovačić & Balaban 2010), word and sentence recognition performance has also been found to improve with younger age at implantation (Fryauf-Bertschy et al. 1997; Harrison et al. 2005; Tong et al. 2007; Uziel et al. 2007) – suggesting a common underlying mechanism for all these abilities.

Broadening the scope of measurement to include processing of both linguistic and indexical information in speech is important, as the perception and subsequent development of spoken language and social skills may be affected by these interwoven processes. Most CI users are enrolled in classes with NH age mates by elementary grades (Geers & Brenner 2003) and may not have acquired these skills that are often learned by passive exposure to events witnessed or overheard in incidental situations (Calderon & Greenberg 2003). For example, perception of indexical cues, particularly those for emotional intent, may contribute to effective social interactions.

The current investigation documents both linguistic and indexical speech perception in children with CIs. Children's perception of words at different intensity levels, sentences in noise, non-words, emotional content, and different talkers are analyzed in relation to child, family, audiological, and educational characteristics, to test the following predictions:

1. If linguistic and indexical abilities are interdependent, then performance will be highly related across all types of speech perception tasks.

2. If linguistic perception depends more on spectral resolution than does indexical perception, then phoneme, word and sentence recognition will be more highly correlated with CI characteristics (e.g., technology, aided thresholds) than talker or emotion perception.

3. If indexical perception is associated with auditory input during a critical period, then higher scores on indexical tasks will be associated with younger age at CI and perhaps better pre-implant hearing.

4. If spoken language acquisition is a product of both linguistic and indexical skills, then language test scores will be related to both types of perceptual measures.

5. If social interaction is facilitated by perception of indexical cues in the speech signal, then social skills ratings will be more highly related to performance on indexical than linguistic measures.

6. Strong speech perception skills, particularly in challenging listening conditions, will be associated with earlier mainstreaming with hearing age-mates.

METHODS

Children with CIs

The sample for this study was recruited from 76 participants in a previous study of children who had received a single CI between 12–38 months of age, and who had been enrolled in listening and spoken language programs since receiving their first CI (Nicholas & Geers 2006, 2007). At the time of initial enrollment in the previous study, all children were 3.5 years of age. All hearing losses were presumed to be congenital. Additional inclusion

criteria were English as the primary language used in the home and normal nonlinguistic developmental milestones based on direct testing or parent interview at age 3.5. Children meeting the inclusion criteria had been recruited from preschool programs and auditory-verbal practices across North America. Families of all participants in the previous study were invited to attend a 2-day research camp in St. Louis for the present study when they were between the ages of 9 and 12 years and 60 families accepted. The etiology of the hearing loss in most cases (N=31) was unknown, 26 were genetic, and a single case each was attributed to ototoxic drug exposure, maternal CMV and meningitis (at age 5 days). All children used hearing aids (HA) prior to cochlear implantation, although most discontinued this use following receipt of a CI.

Child, family, audiological and educational characteristics are summarized in Table 1. Throughout this manuscript the CI group will be referred to as "10-year olds", corresponding to the mean chronologic age of 10.5. Children were administered the nonverbal subtests of the Wechsler Intelligence Scale for Children – 4th Edition (WISC-IV; Wechsler 2003) and the Perceptual Reasoning Quotient (PRQ) of this test was used to estimate nonverbal learning ability. Most (73%) of the children scored within 1 *SD* of the normative average (*M* = 100; *SD* = 15) with 4 below and 12 above average. Participant's families were well-educated (average highest parent education level was a college degree) and had a combined annual income of > $95,000. A composite variable (Parent Education + Family Income category) was created to reflect socioeconomic status (SES).

The Core Language Index of the *Clinical Evaluation of Language Fundamentals, 4th Edition* (CELF-4; Semel et al. 2003) was used to estimate language level across the following subtests: Concepts and Directions (understanding oral commands of increasing length and complexity), Recalling Sentences (repeating increasingly complex sentences), Formulated Sentences (constructing a sentence using stimulus words) and Word Classes (understanding and expressing semantic relations between words). Standard scores expressed each child's performance in relation to age-mates with NH (Mean = 100; SD = 15) and the average CI participant scored within one standard deviation of this average.

The *Social Skills Rating System* (SSRS; Gresham & Elliott 1990) provided a parent's assessment of their child's ability to interact effectively with others and avoid socially unacceptable responses. The component scales (cooperation, assertion, self-control and responsibility) document the child's development of social competence, peer acceptance and adaptive social functioning at school and at home. The average rating for the CI participants (104) was well within the average range for NH age-mates (M = 100; SD = 15).

Historical and current audiological information was obtained from a combination of records review, parent report and direct assessment. Before receiving a CI, the group unaided pure-tone average (PTA) (500, 1k, 2k Hz) in the better ear was 108 dB HL and the average aided PTA was 65 dB HL. The group's current average PTA using their typical device configuration was 21.5 dB, HL. The reported mean age at hearing loss diagnosis was 10 months and mean age provision of HAs was 11 months. Children received their first CI between 12 – 38 months of age (M = 22 months). About half of the sample (N = 29) received a second CI at an average age of 7;8 (years; months) and used bilateral CIs for an average duration of 2;11 at time of test. A Bilateral-CI Rating was assigned to each participant based on duration of use of a 2nd CI: '0' for those without a 2nd CI (N= 31), '1' for < 1 year (N= 14), and '2' for 1 year (N = 15). To examine the effects of processor upgrades on perception, a Technology Rating value was assigned to each processor based on its manufacturer's technology generation. For example, processors manufactured by Cochlear Corporation were rated from earliest to most recent as follows: 1-Spectra, 2-ESPrit 22, 3- Sprint or ESPrit 3G, and 4-Freedom. Processors by Advanced Bionics were

rated as follows: 1-PSP, 2- Platinum BTE, 3- Auria BTE, and 4- Harmony BTE. Fifty-five of the 60 children received an upgrade to newer processor technology in at least one ear between initial implantation and the current test session. Twenty-seven of the 29 children who received a second CI used most-recent processors in one or both ears. Only 15 of the 31 unilateral CI users used most-recent processors.

Educational characteristics of the CI students were also examined. Upon enrolling in the study at 3.5 years of age, most (79%) were enrolled in oral preschool programs; 13% in mainstream preschools and 8% were at home with individual therapy. At the time of the present study (average age = 10; 5), most (90%) were placed in a full-day mainstream classroom with NH age-mates. School grades completed the previous year were 3rd (N = 15), 4th (N = 33), 5th (N = 10) and 6th (N = 2). Children were mainstreamed, on average, by kindergarten and spent almost half (47%) of their elementary years in classes with more than 20 students.

Children with Normal Hearing

Thirty children with normal hearing were recruited from the local community to serve as a comparison group when published data on normally-hearing children were not available. This comparison group had equal numbers (15) of boys and girls. The CI and NH groups did not differ significantly in age (M = 10.4 years; SD = .80) and median family income category ($90,000+). However the NH group had a greater percentage of mothers (57%) completing post graduate work than mothers of the CI group (19%). All NH participants passed a hearing screening at 15 dB HL for octave frequencies from 250–4000 Hz.

Method of Data Collection

CI participants attended one of six identical "research camps" offered in Saint Louis, Missouri on the campus of Washington University School of Medicine during the summers of 2008, 2009, and 2010. The study paid expenses for the participant and one parent. The protocol for the study was approved by the Human Research Protection Office at Washington University in Saint Louis.

Children were administered hearing, speech perception, speech production, language, intelligence, and reading tests. Language outcomes have been reported elsewhere (Geers & Nicholas, in press), indicating that higher language scores in this group of 60 children were associated with younger age at CI and better pre-implant aided hearing. This paper focuses on the speech perception outcomes. Aided sound-field detection thresholds were obtained using frequency modulated (FM) tones at octave frequencies from 250–4000 Hz. Participants were seated approximately 1 to 1.5 m from the loudspeaker at 0° azimuth using their typical device configuration. Mean CI-aided PTA was 21.5 dB HL (SD = 7.3; range 8–48; tested with contralateral HA for N=2). All tests were administered using pre-recorded stimuli presented from a computer. All tests except the non-word repetition task were routed through a GSI 61 audiometer and sound-field loudspeaker. Non-word stimuli were presented via a desktop speaker (Cyber Acoustics MMS-1). All stimuli were calibrated using a Type 2 sound level meter placed at the approximate location of the child's implant microphone.

Linguistic Speech Perception

LNT (Kirk et al. 1995)—Open-set word recognition was tested using monosyllabic, 50-word lists drawn from the vocabulary of 3–5 year old typically-developing children. Each child heard two 50-word lists presented in quiet, one list at a loud conversational level (70 dB SPL) and a second list at a soft level (50 dB SPL). The child was instructed to repeat what he/she heard. For each level-condition, a percent-correct word score was computed that represents the percent of words (responses) that were recognizable as the target word.

School-aged children with NH were found to score > 90% correct on the LNT even at soft levels (Eisenberg et al. 2002). Each word list was comprised of 25 "Lexically Easy" and 25 "Lexically Hard" words based on the frequency of occurrence of the words in the language and density of words within the lexical neighborhoods, as measured by one phoneme substitution. To the extent that linguistic knowledge affects performance, scores are expected to be higher on lexically easy than on lexically hard word-lists and this difference may be greater at softer levels.

Bamford-Kowal-Bench Speech in Noise (BKB-SIN) Test (Killion et al. 2004)— Recorded BKB-SIN sentence list pair #12 (A & B) (Bamford & Wilson 1979) was presented at 65 dB SPL in the presence of four-talker babble; speech and noise were presented at 0° azimuth. For each sentence presentation, the Signal-to-Noise Ratio (SNR) is automatically decreased by 3 dB, starting with a SNR of +21 dB and ending at 0 dB. The SNR for 50% correct was computed for each list (A & B) in the pair and then averaged for a single 'SNR-50' per participant. The mean SNR-50 for NH 7–10 year olds, reported in the BKB-SIN test manual is 0.8 dB; SD = 1.2 (Killion et al. 2004).

Non-word Repetition Test— A shortened, adapted version of the Children's Nonword Repetition test (CNRep; Gathercole & Baddeley 1996) was used to assess perception of speech sounds and suprasegmental properties. The test consists of one list of 20 non-words, 5 each at 2, 3, 4 and 5 syllables, recorded by a female talker at the Indiana University Speech Research Laboratory (Carlson et al. 1998). Digital recordings of the non-words were played in random order at 65 dB SPL. Children were told they would hear a "funny word" and were instructed to repeat it back as best they could. Their imitation responses were recorded via a head-mounted microphone. Later, responses were transcribed by graduate students/clinicians in speech-language pathology using the International Phonetic Alphabet (International Phonetic Association 1999) and scored for both segmental (Dillon et al. 2004) and suprasegmental (Carter et al. 2002) accuracy. Using Computer Assisted Speech and Language Assessment (CASALA) software (Serry & Blamey 1999), four scores were obtained for each participant: (i) overall accuracy (all segmental and suprasegmental features of the target nonword must be repeated correctly), (ii) percent-correct consonants (a consonant must be repeated correctly in the correct position), (iii) percent-correct number of syllables (the repeated non-word must have the same number of syllables as the presented non-word), and (iv) percent-correct syllable stress (the repeated non-word must have the same stress pattern as the presented non-word). Scores of the CI participants were compared with those of the 30 NH children tested as part of this study.

Indexical Speech Perception

These closed-set tasks were administered to CI and NH participants using pre-recorded spoken sentences presented at 60 dB SPL and using the APEX 3 program developed at ExpORL (Francart et al., 2008; Laneau et al., 2005).

Emotion Perception— Perception of a speaker's emotion was assessed using sentences spoken by a female talker with four emotional contents: Angry, Scared, Happy, and Sad. Three semantically-neutral sentences with simple vocabulary were recorded multiple times with each emotion (Uchanski et al. 2009). In the *Emotion Discrimination* task, 24 pairs of sentences were presented and the child made a touch-screen response indicating whether the sentences were spoken with the same feeling or with different feelings. In the *Emotion Identification* task, a total of 36 sentences were presented and after each sentence presentation the child chose among the four emotions represented by faces displayed on a computer screen.

Talker Discrimination—Two tasks required children to discriminate talkers based on voice characteristics. Stimuli were pre-recorded Harvard IEEE sentences (IEEE 1969) and a subset of the Indiana Multi-Talker Speech Database (Karl & Pisoni 1994; Bradlow et al. 1996). In both tasks, the subject always heard two different sentences and indicated whether those sentences were spoken by the same talker or by different talkers. The *Across-Gender Talker Discrimination* task consisted of a mix of 8 male and 8 female talkers' sentence productions. If a trial presented two different talkers, it was always a male vs. female contrast. The *Within-Female Talker Discrimination* task presented a mix of 8 female talkers. If a trial presented two different talkers, then two different female talkers spoke the two sentences. A total of 32 sentence pairs were presented in each condition.

RESULTS

Descriptive results for the test battery are provided in Table 2.

Linguistic Speech Perception

LNT—The majority of CI users scored below the 90–100% level previously documented for NH school-aged children at both soft and loud levels (Eisenberg 2002). Fourteen CI users did score 90% or higher at 70 dB SPL and 2 did so at 50 dB SPL, indicating some children with early CIs performed equivalent to their age-mates with NH. The average LNT word score decreased, from 77% to 52% correct with a 20 dB decrease in signal level, and this difference was statistically significant ($F[23, 36] = 5.31$, $p < .001$). LNT scores at both presentation levels, for each participant, are shown in Figure 1. While some subjects (N=13) obtained similar word scores at both levels (i.e., within 10 percentage points), most children (N = 47) showed a decrease (= 10 percentage points) when the level was reduced from loud to soft. The average difference (LNT 70 score – LNT 50 score) was 24 percentage points, and these differences in LNT scores are correlated negatively with both CI-aided PTA threshold ($r = −.39$; $p = .01$) and the rating of speech processor technology ($r = −.34$; $p = .01$). That is, children who used newer technology and obtained lower sound field thresholds with their CI(s) showed smaller decrements in speech recognition scores when the sound level was reduced.

Mean scores from the lexically-easy LNT word lists were slightly greater than scores from the lexically-hard word lists (79% vs. 74% at 70 dB SPL and 54% vs. 50% at 50 dB SPL). A 2-way ANOVA examined presentation level by list difficulty and found a significant effect for presentation level ($F[1, 236] = 66.79$; $p < .001$), but not for lexical difficulty ($F[1, 236] = 2.40$; $p = .122$). There was no significant interaction between presentation level and list difficulty ($F[1, 236] = .05$; $p = .825$).

BKB-SIN—The average SNR-50 for children with CIs (9.8 dB) was substantially higher than the SNR-50 previously reported for NH 7–10 year olds (0.8 dB) (Killion et al. 2004). The average CI user in this sample needed a 9 dB greater SNR than the average NH 7–10 year old child for an equivalent level of performance. There was a significant advantage for the 29 children who received a second CI (mean SNR-50 = 8.29; SD = 3.48) over those 31 who remained unilateral users (mean SNR-50 = 11.48; SD = 5.15) ($F[1,58] = 7.76$; $p = .01$).

Non-Word Repetition—Mean overall accuracy score for CI users (27% correct) was substantially lower than that observed for the NH group (Mean = 90% correct; SD = 7.9). Though substantially poorer than NH children's scores, these CI participants' scores are considerably higher than those reported by Carter et al. (2002) in which only 5% of non-word imitations by 8–9 year old CI-users were produced without errors. In addition, three CI users scored at or above the minimum NH score of 70% correct, suggesting that

performance equivalent to age-mates with NH is possible for some children with early cochlear implantation. Mean overall accuracy decreased with increasing syllable length, from 33% correct overall production of 2-syllable non-words to 16% correct overall production of 5-syllable non-words.

The average consonant score was 73% correct for the CI group, substantially below the mean of 99% correct obtained for the 30 NH age-mates. However, the mean consonant score of CI users is much greater than the 39% mean consonant score reported by Dillon et al. (2004) for 24 children with CIs. Similar to overall accuracy, consonant performance decreased significantly with syllable-length, from 73% correct for 2-syllable non-words to 63% correct for 5-syllable non-words (t[59] = 4.45; p < .001). The number of consonants correctly imitated in non-words and LNT word recognition scores at 70 dB SPL are highly correlated (r = .829), indicating that similar perceptual abilities may be tapped by real-word and non-word stimuli.

Suprasegmental performance on the non-word repetition task was also examined. NH children repeated the correct number of syllables and correct stress at equivalently high levels (99% and 96%, respectively), while CI users repeated the correct number of syllables with greater accuracy (85% correct) than the proper stress pattern (65% correct) (F[12, 47] = 2.6; p = .009).

Indexical Speech Perception

Performance of the children with CIs on the four indexical tasks was compared with (a) mean chance levels of performance, (b) the upper limit of the 95% confidence interval for chance performance, and (c) the performance of NH age-mates (see captions for Figures 2–5).

Emotion Perception—Table 2 presents average scores on the Emotion Discrimination task, in which participants responded whether two sentences were spoken with the same or different emotions, and on the Emotion Identification task, in which they chose one out of four emotions expressed in a spoken sentence. The mean discrimination score for the CI group (90% correct) was 40 percentage points above mean chance performance, and the mean identification score (58% correct) was 33 percentage points above mean chance performance.

Individual results on the Emotion Discrimination and Emotion Identification tasks are displayed in Figures 2 and 3, respectively. Figure 2 illustrates that fifty-six of the 60 CI participants (93%) performed above 71%, the 95% confidence limit on the discrimination task. Figure 3 indicates that 50 participants (83%) scored above 39%, the 95% confidence limit, on the identification task. Only one child scored below the 95% confidence limit on both the identification and discrimination tasks. Scores exhibited by the 30 NH age-matched are also represented in these figures. Mean scores for the NH group were close to ceiling, at 99% (SD = 2.5) correct emotion discrimination and 88% (SD = 9.0) correct emotion identification. Twenty-two of the children with CIs scored within the range of normal performers (i.e., >92%) on the discrimination task and 17 scored in this range (i.e., above 70%) on the identification task.

Emotion identification results were analyzed further for accuracy for each emotion. For the NH cohort, though overall identification accuracy was 88%, correct identification ranged from 77% for 'scared' to 97% for 'sad,' with intermediate accuracies of 84% for 'angry' and 94% for 'happy.' The children with CIs exhibited a somewhat similar pattern of accuracies for these four emotions, with again the lowest for 'scared' (34%) and then somewhat similar and higher accuracies for 'angry,' 'happy,' and 'sad' (68%, 54% and 67%). Discrimination

of emotion in sentence pairs was also analyzed further for same vs. different trials. The NH cohort performed nearly perfectly (98–100%) for all same-pairings (e.g., 'angry vs. angry', 'sad vs. sad,' etc.) and all different-contrasts ('angry vs. sad', 'happy vs. sad,' etc.), except for 'sad vs. sad' which had a slightly lower accuracy of 94%. The children with CIs correctly identified 'same' emotions as 'same' 94% of the time and 'different' emotions as 'different' 86% of the time. Correct responses for same-pairings were similar for all emotions, 91–97%, while different-contrasts varied in accuracy from 71% for 'scared vs. happy' to 96% for 'angry vs. sad.'

Perception of Talker Differences—Children with CIs averaged close to 90% correct on the discrimination task that paired male vs. female talkers. As a group, their performance was only slightly below that of NH age-mates, who averaged 98% correct for this task. As might be expected, the task of discriminating between two female talkers was more difficult; the CI group average was 65% correct, which is substantially below the NH average of 91% correct. Individual results for talker discrimination are presented in Figures 4 and 5. Most of the children could discriminate talkers to some extent; 90% (N=54) and 58% (N=35) scored significantly above chance on the Across-Gender and Within-Female tasks, respectively.

Linguistic and Indexical Speech Perception Composite Scores

Table 3 presents inter-correlations among the linguistic measures. A principal components (PC) analysis of these scores revealed that all tests loaded highly (i.e., all factor loadings > 0.8) on a single factor that accounts for 79% of common variance (Eigenvalue = 3.948). Combining these measures into a single composite variable score reduces the error associated with each individual test score (Strube 2003). Higher PC loadings on some measures (e.g., BKB-SIN, NWR cons) reflect their higher relation to the new composite variable. The composite PC score is used as the outcome measure (dependent variable) in all further analyses, and will be referred to as 'Linguistic Speech Perception (LSP).'

A similar analysis was conducted for the four indexical measures and components of the "Indexical Perception" score are summarized in Table 4. Here the loadings were not quite as high as for the Linguistic components, but a single factor was identified that accounted for 62% of common variance, with individual loadings ranging between .73 and .84 and an Eigenvalue of 2.467. This composite score will be referred to as 'Indexical Speech Perception (ISP)'. The LSP Score and ISP Scores were highly correlated (r =.764; p<.001), suggesting that all of these measures share common variance, and ISP and LSP are interdependent. However separate analyses of LSP and ISP scores were conducted to determine the extent to which they are associated with different underlying characteristics.

Associations with Predictor Variables

Correlates of LSP and ISP scores were examined for: 1) child and family characteristics; 2) audiological characteristics; 3) educational program characteristics. Results are summarized in Table 5.

Child and Family Characteristics—Neither LSP nor ISP scores were significantly related to the child's nonverbal intelligence (the Perceptual Reasoning quotient on the WISC-IV) or to the family SES estimate (parent education and family income). As predicted, well-developed LSP and ISP skills were associated with language scores closer to NH age-mates. However, only ISP skills were significantly associated with social skills ratings, suggesting an important role of talker identification and recognition of talker emotions in social development. This novel finding must be interpreted cautiously, due to insufficient power for direct statistical comparison of linguistic and indexical correlations with social skills ratings. However, such a difference could demonstrate an important

dissociation between these two parallel sources of information encoded in the speech signal. Social-emotional development may be especially vulnerable in children who have difficulty encoding and processing indexical properties of speech signals related to vocal source information of different talkers.

Audiological Characteristics—Younger age at implantation was associated with higher ISP and LSP scores, but the relation for ISP skills was stronger, suggesting a critical role of early auditory experience for recognizing voice characteristics. In contrast to this, children with better aided hearing before cochlear implantation had the best linguistic speech perception skills following cochlear implantation but not necessarily the best indexical skills. Both LSP and ISP scores were higher for children who used the most recent CI processor technology, those who achieved lower CI-aided thresholds, and those who received a second, bilateral CI. To further examine the apparent advantage of bilateral CIs for speech perception, three separate contrasts were tested statistically using one-way ANOVA, each for the LSP and ISP: (1) bilateral [N=29] vs unilateral [N=31] users, (2) most-recent [N=42] vs. older [N=18] processor technology users, and (3) within the group of most-recent technology users, bilateral [N=27] vs unilateral [N=15] users. Bilateral CI users performed significantly better than unilateral CI users (LSP: $F(1, 58) = 8.76$, $p < .01$; ISP: $F(1, 58) = 9.27$, $p < .01$). Users of the most-recent technology processors performed significantly better than users of older technology processors (LSP: $F(1,58) = 13.29$, $p < .001$; ISP: $F(1,58) = 20.2$, $p < .001$). However, within the group of listeners using most-recent technology processors, there were no significant differences between the bilateral and unilateral implant recipients. These results therefore suggest that the better performance on both ISP and LSP measures for bilateral users was associated with those participants' use of newer CI processors.

Educational Characteristics—Younger age at mainstreaming was associated with better ISP and LSP scores, indicating that both types of speech perception skills were associated with leaving special education earlier. However, only LSP was related to the percentage of elementary school grades enrolled in large classes (> 20 students). This finding probably reflects the importance of linguistic speech perception skills, particularly in challenging listening conditions, for academic success in typical mainstream classrooms.

DISCUSSION

This study described linguistic and indexical speech perception in sixty 9–12 year olds with congenital hearing loss who received early listening and spoken language intervention, and a unilateral CI by 38 months of age. The analyses addressed the following hypotheses:

1. If linguistic and indexical abilities are interdependent, then performance will be highly related across all types of speech perception tasks

Scores on this battery of speech perception tests with a range of test formats and stimuli (open-set, closed-set; non-word, word, sentence) were highly related to one another, reflecting a unified perceptual ability. The strong relation between linguistic and indexical speech perception measures ($r = .76$) is consistent with recent findings with infants (Johnson et al. 2011), adults (Perrachione et al. 2009), adults with dyslexia (Perrachione et al. 2011) and adults who use CIs (Li & Fu 2011). Results strongly support findings summarized by Pisoni (1997) of a very close connection between the form and content of the linguistic message and suggest common underlying processes and representations, such as short-term memory skills and long-term categories for voices. The time-course of development of these cognitive processes and categories are not yet understood in children with impaired or with

normal hearing (e.g., see Sidtis & Kreiman 2012, for a discussion of familiar voice recognition vs non-familiar voice discrimination), and warrant further research.

2. If linguistic perception depends more on spectral resolution than does indexical perception, then phoneme, word and sentence recognition will be more highly correlated with audiological characteristics (e.g., technology, aided thresholds) than talker or emotion perception

Contrary to this prediction, both LSP and ISP scores correlated significantly with audiological characteristics, although the correlations were consistently higher for LSP scores. There was a substantial benefit for both skills associated with more recent speech processor technology. The apparent advantage of bilateral cochlear implantation for both types of speech perception appears to be associated with use of newer technology rather than use of two devices. CI processors associated with a 2nd CI (implanted sequentially, perhaps several years after a 1st CI) will often be newer technology CI processors that incorporate many features which can improve both linguistic and indexical speech perception.

3. If indexical perception is associated with auditory input during a critical period of development, then higher scores on indexical tasks will be associated with younger age at CI and perhaps better pre-implant hearing

This hypothesis was confirmed by a significant association between earlier implantation and better indexical skill development. However, this relation was not observed for pre-implant hearing level, which was correlated with LSP rather than ISP scores. This apparently contradictory finding may be associated with the level of profound deafness exhibited by the vast majority of these children pre-implant. Early acoustic hearing may have resulted in some linguistic benefit, but it was the strong auditory signal provided by the CI at a young age that helped these children to learn discrimination of vocal characteristics.

4. If spoken language acquisition is a product of both linguistic and indexical skills, then language test scores will be related to both types of perceptual measures

This hypothesis was confirmed, with strong correlations between both LSP and ISP skills and language test scores. This relation appeared to be specific to language and was not associated with significant relations to nonverbal cognitive development. On the other hand, scores on the *LNT* were not affected significantly by lexical difficulty, as might be anticipated if language skills influenced speech perception scores. The current LNT results were compared with LNT scores from the large sample (N=181) of 8- and 9-year-old children with CIs reported by Geers et al. (2003) where there was a significant effect of lexical difficulty. Mean scores of the current sample (at 70 dB SPL) were ~30 percentage points higher (79% compared to 48% for easy word lists and 74% compared to 44% for hard word lists), suggesting that improved speech perception in the current sample, possibly resulting from younger age at implant, more recent CI technology and enrollment in an exclusively spoken language environment, may have restricted the influence of lexical neighborhood density on word recognition scores.

5. If social interaction is facilitated by perception of indexical cues in the speech signal, then social skills ratings will be more highly related to performance on indexical than linguistic measures

This hypothesis was confirmed by a significant correlation between ISP scores and social skills ratings. This relation was not found for LSP skills. Well-developed social skills were more highly associated with the ability to discriminate the nuances of talker identity and emotion than with the ability to recognize words and sentences through listening. This is consistent with Schorr et al. (2009), where quality of life ratings for CI users 5 to 14 years

old were found to be associated with their performance on an emotion discrimination task but not on a word recognition test (the LNT).

6. Strong linguistic speech perception skills, particularly in challenging listening conditions, will be associated with earlier mainstreaming with hearing age-mates

This result was confirmed by our finding of significant correlations between both ISP and LSP scores and the age at which children entered mainstream classes. This means that children with better speech perception demonstrated the potential for successful academic integration sooner than those with poorer auditory perception particularly when listening in noise and understanding speech at a soft level.

Better linguistic perception skills were also associated with placement in more typically-sized classrooms (i.e., >20 students) for a larger proportion of their elementary school experience. Comprehension in larger classes would certainly be facilitated by better speech perception skills in degraded listening conditions, such as those measured in this study. These linguistic speech perception results highlight the vulnerability of this population to poor listening environments when compared to their NH age-mates. For example, the mean SNR necessary to understand 50% of key words in sentences was roughly 10 dB, about 9 dB higher than needed for NH age-mates. The average SNR across a variety of classroom locations is only about 11 dB (Sato & Bradley 2008). Hence, for many of these children, only a portion of classroom speech is intelligible through listening alone, though it is clearly intelligible to their NH classmates. Some of this disadvantage can be overcome through use of FM or sound field systems. When asked about such systems, 48 (of our 60) participants reported they use them in the classroom either some or all of the time. However 12 students reported never using FM systems and thus may be missing significant amounts of speech information in school.

CONCLUSIONS

Most participants with CIs correctly recognized both linguistic and indexical content in speech, with average scores above 75% for open-set recognition of words and significantly above chance on closed-set tasks measuring both discrimination and identification of talker characteristics. However scores of most CI users did not reach NH levels on any of the tasks administered and performance on the within-female talker discrimination task was especially difficult for these children.

Linguistic Speech Perception

For LSP skills, there was a ~ 25 percentage point reduction in mean word recognition scores with a 20 dB decrease in intensity level, and an average ~ 10 dB greater speech relative to noise level was required for understanding sentences. While average segmental and suprasegmental scores of the CI users on the non-word imitation task approached NH levels (73% and 75% correct compared to 99% and 98% in NH age-mates), accurate reproduction of entire non-words was much lower (mean = 27% compared to 99% in NH). However, performance of these CI users was substantially above non-word scores reported for CI users in earlier studies (Carter et al. 2002; Cleary et al. 2002), possibly due to younger age at implantation and more recent CI technology.

Indexical Speech Perception

The majority of the CI listeners (63%) scored within the range of NH listeners on the Emotion Discrimination task but only 30% scored in the NH range on the Emotion Identification task. This level of performance is consistent with the listeners relying on well-perceived envelope-intensity cues, which are reportedly useful for perceiving emotional

prosody (Chen et al. 2012). Most CI listeners could discriminate male from female talkers fairly easily, but had great difficulty discriminating female talkers from each other. The proportion of CI children scoring within the range of NH age-mates dropped from 65% for the Across-Gender discrimination task to 8.5% for the Within-Female task. This result is consistent with others (Fu et al. 2004, Fu et al. 2005; Cullington & Zeng 2011).

Results of this study support providing a CI as close to 12 months of age as possible, upgrading to the most up-to-date speech processors as they become available, and programming CIs to achieve the lowest possible aided thresholds. These efforts should lead to improved linguistic and indexical speech perception skills. Linguistic and Indexical speech perception skills are highly related and appear to be interdependent, with good skills in one area facilitating performance in the other. These two streams of information appear to be processed together in parallel by the auditory system and central auditory pathways. Facility with LSP and ISP is associated with higher levels of language development and earlier placement in mainstream classrooms. Facility with ISP is associated with higher levels of social skills in the mid-elementary grades. Further research replicating these findings, particularly the correlations of indexical and linguistic skills with social-emotional outcomes, is warranted to verify the significant relation observed between indexical perception and social development.

Acknowledgments

References

Abercrombie, D. Elements of general phonetics. Chicago: Aldine Pub; 1967.

Bachorowski JA, Owren MJ. Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. J Acoust Soc Am. 1999; 106:1054–1063. [PubMed: 10462810]

Bamford, J.; Wilson, I. Methodological considerations and practical aspects of the BKB Sentence Lists. In: Bench, J.; Bamford, J., editors. Speech-hearing tests and the spoken language of hearing-impaired children. London: Academic Press; 1979.

Banse R, Scherer KR. Acoustic profiles in vocal emotion expression. J Pers Soc Psych. 1996; 70:614–636.

Blamey PJ, Sarant JZ, Paatsch LE, et al. Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. J Speech Lang Hear Res. 2001; 44:264–285. [PubMed: 11324650]

Bond M, Elston J, Mealing S, et al. Effectiveness of multi-channel unilateral cochlear implants for profoundly deaf children: a systematic review. Clin Otolaryngol. 2009; 34:199–211. [PubMed: 19531168]

Botting N, Conti-Ramsden G. Non-word repetition and language development in children with specific language impairment (SLI). International Journal of Language and Communication Disorders. 2001; 36:421–432. [PubMed: 11802495]

Bradlow AR, Torretta GM, Pisoni DB. Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. Speech Commun. 1996; 20:255–272. [PubMed: 21461127]

Bradlow AR, Nygaard LC, Pisoni DB. Effects of talker, rate, and amplitude variation on recognition memory for spoken words. Percept Psychophys. 1999; 61:206–219. [PubMed: 10089756]

Calderon, R.; Greenberg, M. Social and Emotional Development of Deaf Children. In: Marshark, M.; Spencer, P., editors. Deaf Studies, Language and Education. Washington DC: Gallaudet University; 2003. p. 177-189.

Carlson, JL.; Cleary, M.; Pison, DB. Research in Spoken Language Processing Progress Report No 22. Bloomington, IN: Speech Research Laboratory, Indiana University; 1998. Performance of normal-hearing children on a new working memory span task; p. 251-273.

Carter A, Dillon C, Pisoni D. Imitation of nonwords by hearing impaired children with cochlear implants: Suprasegmental analylses. Clinical Linguistics & Phonetics. 2002; 16:619–638. [PubMed: 12596429]

Chen X, Yang J, Gan S, et al. The contribution of sound intensity in vocal emotion perception: Behavioral and electrophysiological evidence. PloS ONE. 2012; 7:e30278. [PubMed: 22291928]

Cheng AK, Grant GD, Niparko JK. Meta-analysis of pediatric cochlear implant literature. Ann Otol Rhinol Laryngol Suppl. 1999; 177:124–128. [PubMed: 10214816]

Cleary M, Dillon C, Pisoni DB. Imitation of nonwords by deaf children after cochlear implantation: Preliminary findings. Ann Otol Rhinol Laryngol Suppl. 2002; 189:91–96. [PubMed: 12018358]

Cleary M, Pisoni DB, Kirk KI. Influence of voice similarity on talker discrimination in children with normal hearing and children with cochlear implants. J Speech Lang Hear Res. 2005; 48:204–223. [PubMed: 15938065]

Cleary M, Pisoni D. Taker discrimination by prelingually deaf children with cochlear implants: Preliminary Results. Annals of Otology, Rhinology & Laryngology. 2002; 111(Supp):113–118.

Cosendai G, Pelizzone M. Effects of the acoustical dynamic range on speech recognition with cochlear implants. Audiol. 2001; 40:272–281.

Craik FM, Kirsner K. The effect of speaker's voice on word recognition. Quarterly J Exp Psychol. 1974; 26:274–284.

Cullington HE, Zeng FG. Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. Ear Hear. 2011; 32:16–30. [PubMed: 21178567]

Davidson L, Geers A, Blamey P, et al. Factors contributing to speech perception scores in long-term pediatric cochlear implant users. Ear Hear. 2011; 32:19–26.

Davidson LS. Effects of stimulus level on the speech perception abilities of children using cochlear implants or digital hearing aids. Ear Hear. 2006; 27:493–507. [PubMed: 16957500]

Davidson LS, Geers AE, Brenner C. Cochlear implant characteristics and speech perception skills of adolescents with long-term device use. Otol Neurotol. 2010; 31:1310–1314. [PubMed: 20616759]

Davidson LS, Skinner MW, Holstad BA, et al. The effect of instantaneous input dynamic range setting on the speech perception of children with the nucleus 24 implant. Ear Hear. 2009; 30:340–349. [PubMed: 19322085]

Dawson PW, Decker JA, Psarros CE. Optimizing dynamic range in children using the nucleus cochlear implant. Ear Hear. 2004; 25:230–241. [PubMed: 15179114]

Dawson PW, Vandali AE, Knight MR, et al. Clinical evaluation of expanded input dynamic range in Nucleus cochlear implants. Ear Hear. 2007; 28:163–176. [PubMed: 17496668]

Dillon C, Pisoni DB, Cleary &, Carter AK. Nonword imitation by children with cochlear implants. Arch Otolaryngol Head Neck Surg. 2004; 130:587–591. [PubMed: 15148181]

Donaldson GS, Allen SL. Effects of presentation level on phoneme and sentence recognition in quiet by cochlear implant listeners. Ear Hear. 2003; 24:392–405. [PubMed: 14534410]

Eisenberg L, Martinez AS, Holowecky S, et al. Recognition of lexically controlled words and sentends by children with normal hearing and children with cochlear implants. Ear Hear. 2002; 23:450–461. [PubMed: 12411778]

Firszt JB, Holden LK, Skinner MW, et al. Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. Ear Hear. 2004; 25:375–387. [PubMed: 15292777]

Francart T, vanWieringen A, Wouters J. Apex 3: A multi-purpose test platform for auditory psychophysical experiments. J Neuroscience Methods. 2008:283–293.

Fryauf-Bertschy H, Tyler RS, Kelsay DM, et al. Cochlear implant use by prelingually deafened children: the influences of age at implant and length of device use. J Speech Lang Hear Res. 1997; 40:183–199. [PubMed: 9113869]

Fu QJ, Chinchilla S, Galvin JJ. The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users. J Assoc Res Otolaryngol. 2004; 5:253–260. [PubMed: 15492884]

Fu QJ, Chinchilla S, Nogaki G, et al. Voice gender identification by cochlear implant users: The role of spectral and temporal resolution. J Acoust Soc Am. 2005; 118:1711–1718. [PubMed: 16240829]

Gathercole, S.; Baddeley, A. The Children's Test of Non-Word Repetition. London, England: Psychological Corporation, Europe; 1996.

Geers A, Nicholas J. Enduring advantages of early cochlear implantation for spoken language development. J Speech Lang Hear Res. (in press).

Geers, A.; Brenner, C.; Davidson, L. Speech perception changes in children switching from M-Peak to SPEAK coding strategy. In: Waltzman, CNS., editor. Cochlear Implants. New York: Thieme Publications; 1999. p. 211

Geers A, Brenner C. Background and educational characteristics of prelingually deaf children implanted by five years of age. Ear Hear. 2003; 24(Suppl):2–14.

Geers A, Brenner C, Davidson L. Factors associated with development of speech perception skills in children implanted by age five. Ear Hear. 2003; 24:24S–35S. [PubMed: 12612478]

Goldinger SD, Pisoni DB, Logan JS. On the nature of talker variability effects on recall of spoken word lists. J Exp Psychol Learn Mem Cogn. 1991; 17:152–162. [PubMed: 1826729]

Grant KW, Walden BE. Evaluating the articulation index for auditory-visual consonant recognition. J Acoust Soc Am. 1996; 100:2415–2424. [PubMed: 8865647]

Gresham, FM.; Elliott, SN. The Social Skills Rating System. Circle Pines, MN: American Guidance Service; 1990.

Harrison RV, Gordon KA, Mount RJ. Is there a critical period for cochlear implantation in congenitally deaf children? Analyses of hearing and speech perception performance after implantation. Dev Psychobiol. 2005; 46:252–261. [PubMed: 15772969]

Holden LK, Skinner MW, Fourakis MS, et al. Effect of increased IIDR in the nucleus freedom cochlear implant system. J Am Acad Audiol. 2007; 18:777–793. [PubMed: 18354886]

Hopyan-Misakyan TM, Gordon KA, Dennis M, et al. Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. Child Neuropsychol. 2009; 15:136–146. [PubMed: 18828045]

IEEE SC. IEEE recommended practice for speech quality measurements. IEEE Transactions on Audio and Electroacoustics. 1969; 17:225–246.

International Phonetic Association. Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet. Cambridge, England: Cambridge University Press; 1999.

James CJ, Blamey PJ, Martin L, et al. Adaptive dynamic range optimization for cochlear implants: a preliminary study. Ear Hear. 2002; 23:49S–58S. [PubMed: 11883767]

James CJ, Skinner MW, Martin LF, et al. An investigation of input level range for the nucleus 24 cochlear implant system: Speech perception performance, program preference, and loudness comfort ratings. Ear Hear. 2003; 24:157–174. [PubMed: 12677112]

Johnson EK, Westrek E, Nazzi T, et al. Infant ability to tell voices apart rests on language experience. Dev Sci. 2011; 14:1002–1011. [PubMed: 21884316]

Karl, JR.; Pisoni, DB. Res on Spoken Lang Proc Prog Rep No 19. Bloomington, IN: Indiana Univ; 1994. Effects of stimulus variability on recall of spoken sentences: A first report.

Killion MC, Niquette PA, Gudmundsen GI, et al. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2004; 116:2395–2405. [PubMed: 15532670]

Kirk KI, Pisoni DB, Osberger MJ. Lexical effects on spoken word recognition by pediatric cochlear implant users. Ear Hear. 1995; 16:470–481. [PubMed: 8654902]

Kova i D, Balaban E. Voice gender perception by cochlear implantees. J Acoust Soc Am. 2009; 126:762–775. [PubMed: 19640042]

Kova i D, Balaban E. Hearing history influences voice gender perceptual performance in cochlear implant users. Ear Hear. 2010; 31:806–814. [PubMed: 20683189]

Laneaus J, Boets B, Moomen M, vanWieringen A, Wouters J. A flexible auditory research platform using acoustic or electric stimuli for adults and young children. J Neuroscience Methods. 2005:131–136.

Li T, Fu QJ. Voice gender discrimination provides a measure of more than pitch-related perception in cochlear implant users. Int J Audiol. 2011; 50:498–502. [PubMed: 21696330]

Luo X, Fu QJ, Galvin JJ. Vocal emotion recognition by normal-hearing listeners and cochlear implant users. Trends Amplif. 2007; 11:301–315. [PubMed: 18003871]

McDermott HJ, Henshall KR, McKay CM. Benefits of syllabic input compression for users of cochlear implants. J Am Acad Audiol. 2002; 13:14–24. [PubMed: 11833989]

Most T, Aviner C. Auditory, visual, and auditory-visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. J Deaf Stud Deaf Educ. 2009; 14:449–464. [PubMed: 19398533]

Nakata T, Trehub SE, Kanda Y. Effect of cochlear implants on children's perception and production of speech prosody. J Acoust Soc Am. 2012; 131:1307–1314. [PubMed: 22352504]

Nicholas J, Geers A. Effects of early auditory experience on the spoken language of deaf children at 3 years of age. Ear Hear. 2006; 27:286–98. [PubMed: 16672797]

Nicholas J, Geers A. Will they catch up? The role of age at cochlear implantation in the spoken language development of children with severe-profound hearing loss. J Speech Lang Hear Res. 2007; 50:1048–62. [PubMed: 17675604]

Nowicki S Jr, Duke MP. Individual differences in the nonverbal communication of affect: The diagnostic analysis of nonverbal accuracy scale. J Nonverbal Behav. 1994; 18:9–35.

Perrachione TK, Del Tufo SN, Gabrieli JD. Human voice recognition depends on language ability. Science. 2011; 333:595. [PubMed: 21798942]

Perrachione TK, Pierrehumbert JB, Wong PC. Differential neural contributions to native- and foreign-language talker identification. J Exp Psychol Hum Percept Perform. 2009; 35:1950–1960. [PubMed: 19968445]

Pisoni, D. Some thoughts on "normalization" in speech perception. In: Johnson, D.; Mullennix, JW., editors. Talker variability in speech processing. San Diego, CA: Academic Press; 1997. p. 9-32.

Pisoni DB, Kronenberger WG, Roman AS, Geers AE. Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. Ear Hear. 2011; 32(Suppl):60S–74S. [PubMed: 21832890]

Santarelli R, Magnavita V, Filippi RD, et al. Comparison of speech perception performance between Sprint/Esprit 3G and Freedom processors in children implanted with Nucleus cochlear implants. Otol Neurotol. 2009; 30:304–312. [PubMed: 19225440]

Sato H, Bradley JS. Evaluation of acoustical conditions for speech communication in working elementary school classrooms. J Acoust Soc Am. 2008; 123:2064–2077. [PubMed: 18397014]

Scherer KR. Vocal communication of emotion: A review of research paradigms. Speech Commun. 2003; 40:227–256.

Schorr AE, Roth FP, Fox NA. Quality of life for children with cochlear implants: perceived benefits and problems and the perception of single words and emotional sounds. J Speech Lang Hear Res. 2009; 52:141–152. [PubMed: 18664684]

Schvartz KC, Chatterjee M. Gender identification in younger and older adults: use of spectral and temporal cues in noise-vocoded speech. Ear Hear. 2012; 33:411–420. [PubMed: 22237163]

Semel, E.; Wiig, EH.; Secord, WA. Clinical Evaluation of Language Fundamentals. San Antonio, TX: PsychCorp/Harcourt Assessment; 2003.

Serry TA, Blamey PJ. A 4-year investigation into phonetic inventory development in young cochlear implant users. J Speech Lang Hear Res. 1999; 42:141–154. [PubMed: 10025550]

Sidtis D, Kreiman J. In the beginning was the familiar voice: personally familiar voices in the evolutionary and contemporary biology of communication. Integr Psychol Behav Sci. 2012; 46:146–159. [PubMed: 21710374]

Skinner MW, Clark GM, Whitford LA, et al. Evaluation of a new spectral peak coding strategy for the Nucleus 22 Channel Cochlear Implant System. Am J Otol. 1994; 15(Suppl 2):15–27. [PubMed: 8572106]

Skinner MW, Fourakis MS, Holden TA, et al. Identification of speech by cochlear implant recipients with the Multipeak (MPEAK) and Spectral Peak (SPEAK) speech coding strategies. I. Vowels. Ear Hear. 1996; 17:182–197. [PubMed: 8807261]

Skinner MW, Fourakis MS, Holden TA, et al. Identification of speech by cochlear implant recipients with the multipeak (MPEAK) and spectral peak (SPEAK) speech coding strategies II. Consonants. Ear Hear. 1999; 20:443–460. [PubMed: 10613383]

Spahr AJ, Dorman MF. Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices. Arch of Otolaryngol Head Neck Surg. 2004; 130:624–628. [PubMed: 15148187]

Strube M. Statistical analysis and interpretation in a study of prelingually deaf children implanted before five years of age. Ear Hear. 2003; 24:15S–23S. [PubMed: 12612477]

Tong MCF, Leung EKS, Au A, et al. Age and outcome of cochlear implantation for patients with bilateral congenital deafness in a Cantonese-speaking population. Ear Hear. 2007; 28(Suppl):56–58.

Uchanski RM, Davidson LS, Quadrizius S, et al. Two ears and two (or more?) devices: A pediatric case study of bilateral profound hearing loss. Trends Amplif. 2009; 13:107–123. [PubMed: 19447765]

Uziel AS, Sillon M, Vieu A, et al. Ten-year follow-up of a consecutive series of children with multichannel cochlear implants. Otol Neurotol. 2007; 28:615–628. [PubMed: 17667770]

Wechsler, D. Wechsler Intelligence Scale for Children. 4. San Antonio, TX: PsychCorp/Harcourt Assessment; 2003.

Wilson BS, Lawson DT, Finley CC, et al. Coding strategies for multichannel cochlear prostheses. Am J Otol. 1991; 12:S56–S61.

Zeng FG, Grant G, Niparko J, et al. Speech dynamic range and its effect on cochlear implant performance. J Acous Soc Am. 2002; 111:377–386.
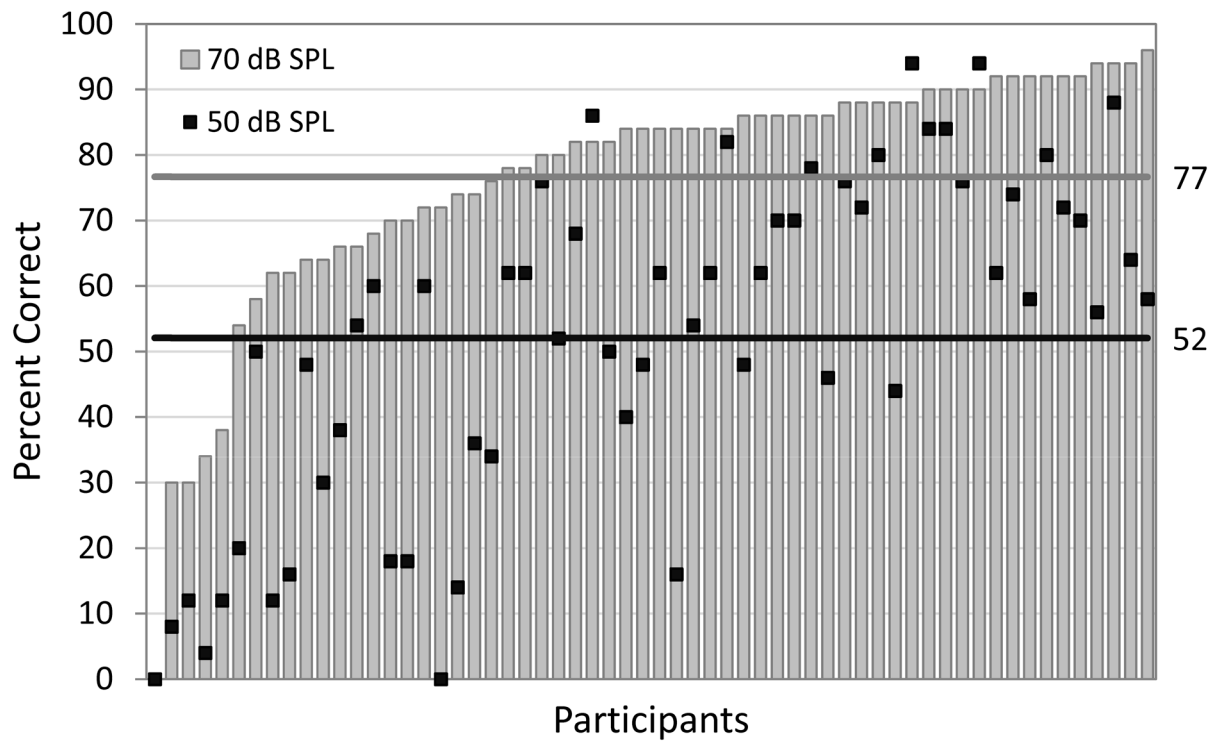
Figure 1.
Percent correct scores on the Lexical Neighborhood Test are plotted for each of 60 children with cochlear implants. Data points represent scores at a presentation level of 50 dB SPL (mean 52%) and columns represent scores at 70 dB SPL (mean 77%). Subjects are ordered by their percent correct at 70 dB SPL scores.

Figure 2.
Columns represent individual percent correct scores for 90 children (60 with cochlear implants and 30 normal hearing) on a same/different emotion discrimination task. Horizontal lines represent chance performance (solid at 50%) and the 95% confidence limit (dashed at 71%).

Figure 3.
Columns represent individual percent correct scores for 90 children (60 with cochlear implants and 30 normal hearing) on a 4-choice emotion identification task (Angry, Scared, Happy and Sad). Horizontal lines represent chance performance (solid at 25%) and the 95% confidence limit (dashed at 39%).
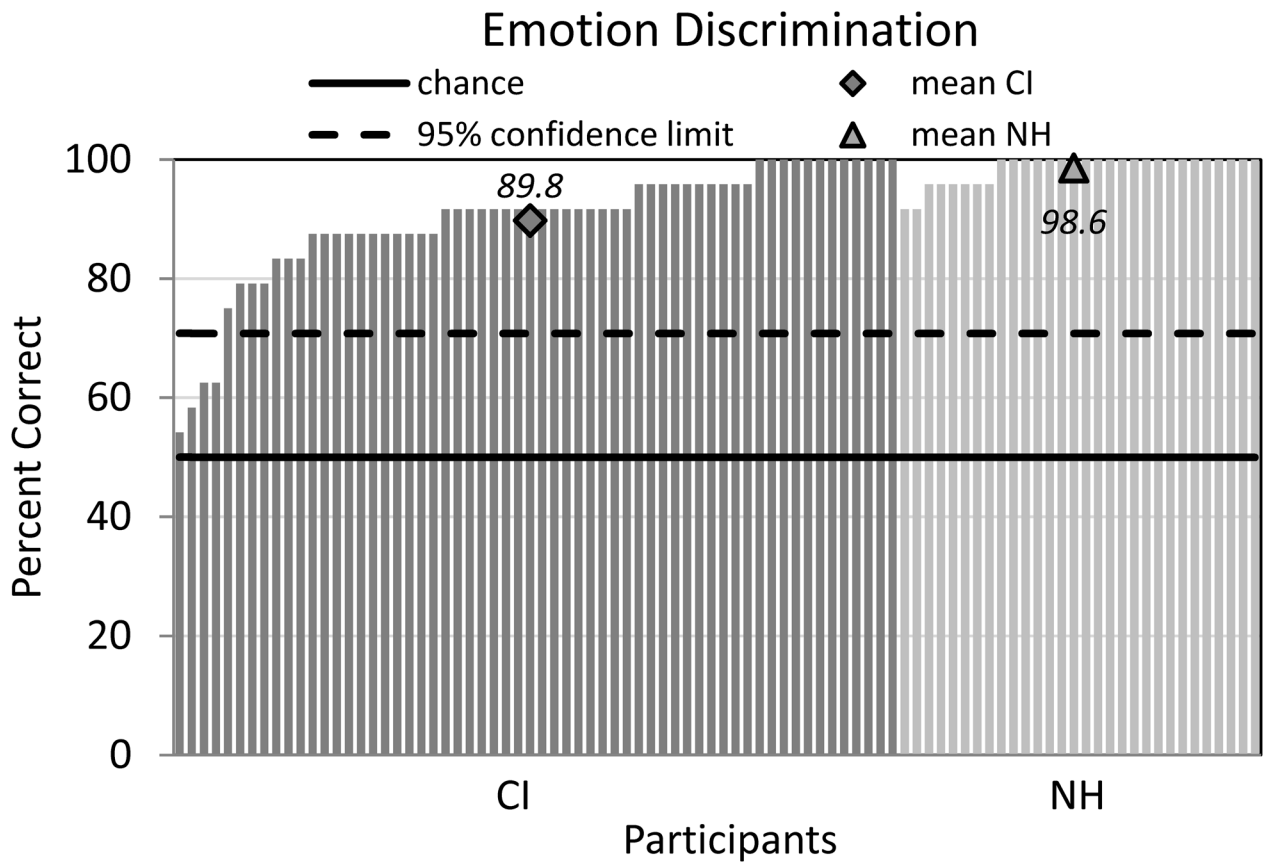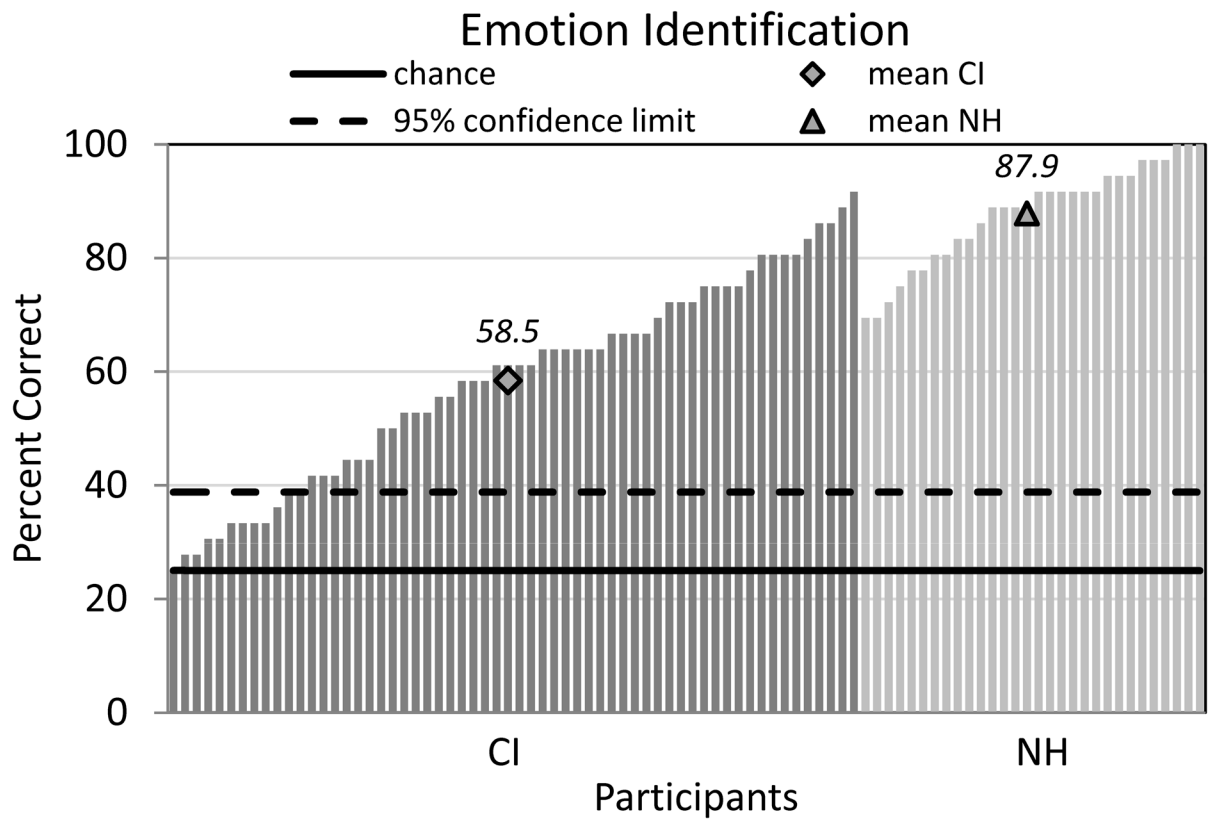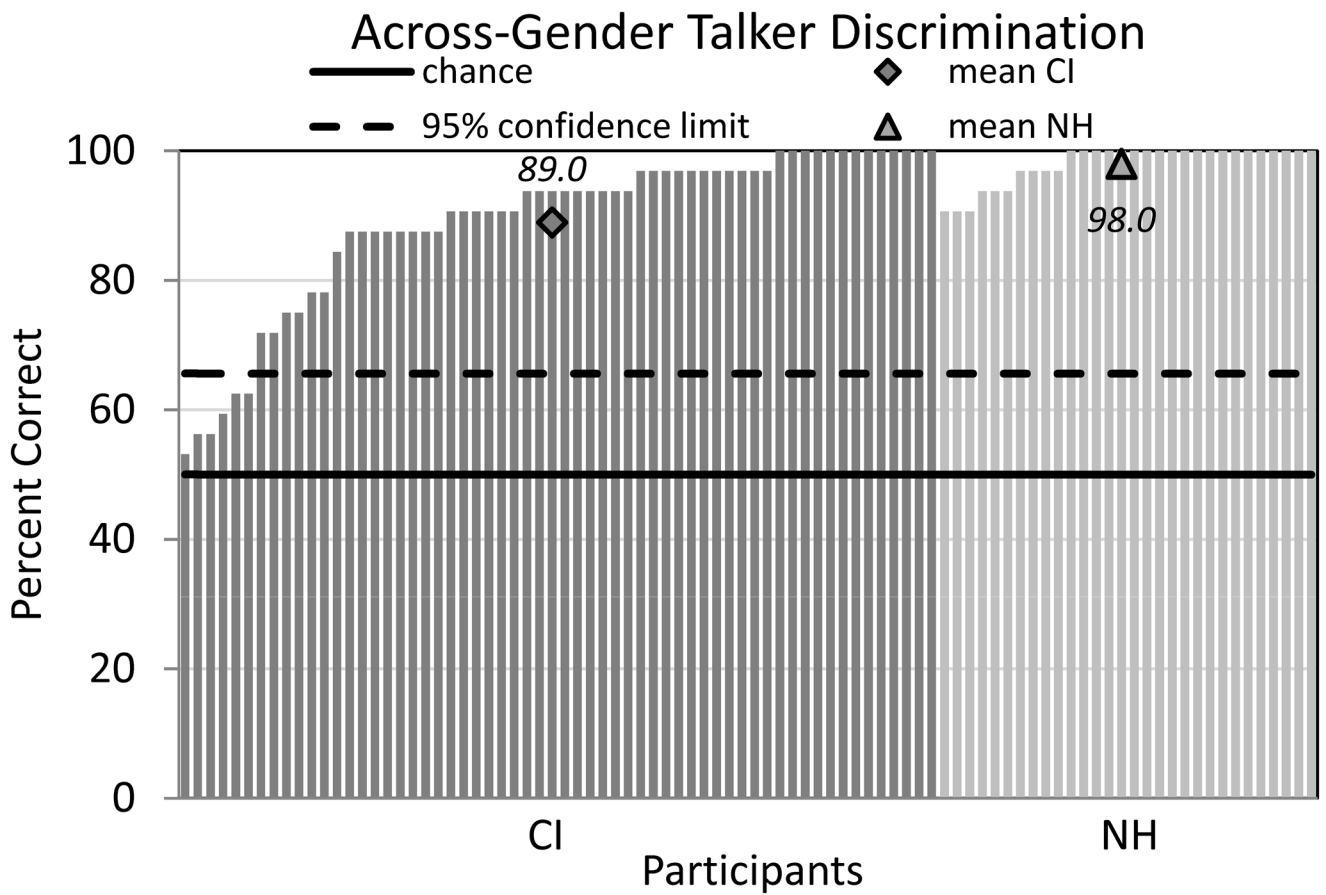
Figure 4.
Columns represent individual percent correct scores for 90 children (60 with cochlear implants and 30 normal hearing) on a same/different Across-Gender talker discrimination task (Male vs Female). Horizontal lines represent chance performance (solid at 50%) and the 95% confidence limit (dashed at 66%).
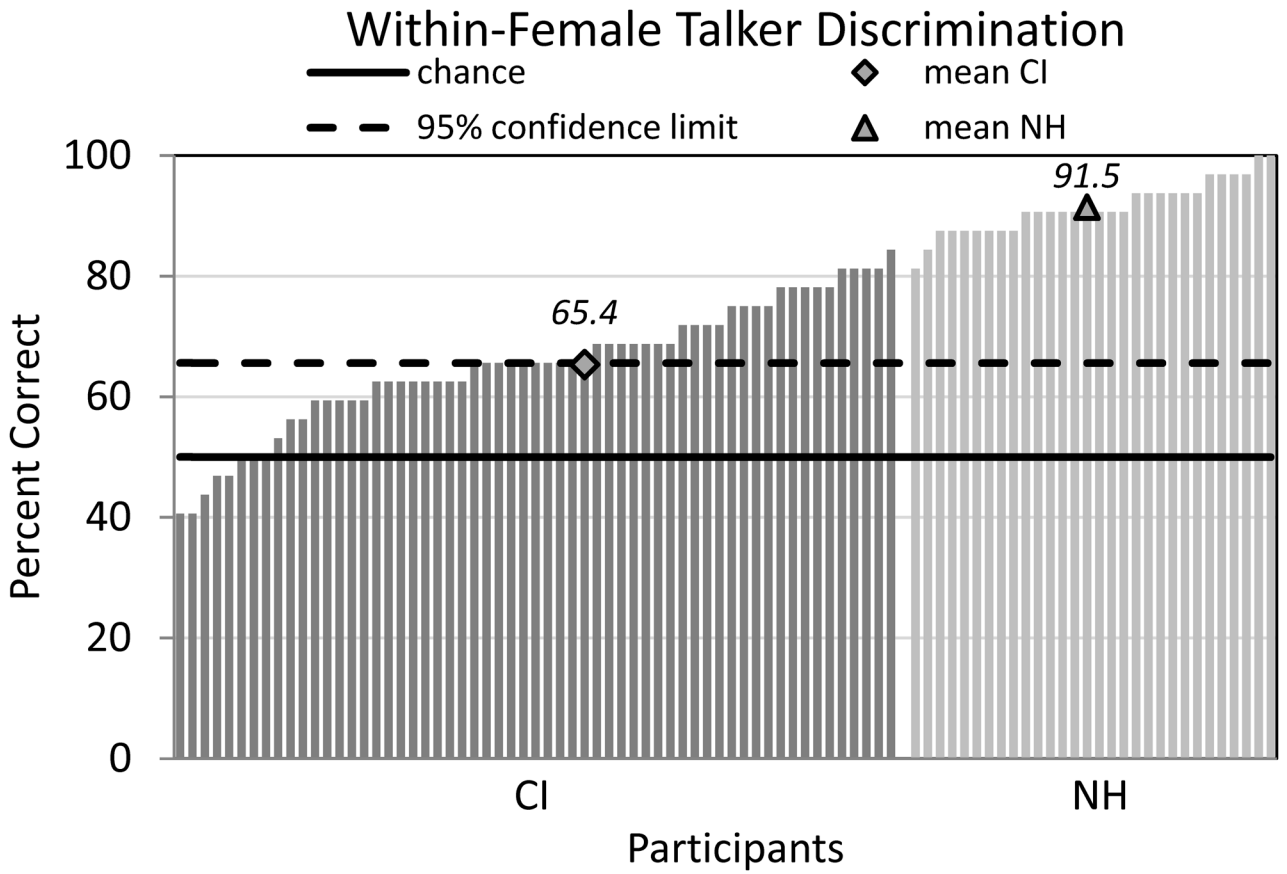
Figure 5.
Columns represent individual percent correct scores for 89 of 90 children (59 with cochlear implants and 30 normal hearing) on a same/different Within-Gender talker discrimination task (Female vs Female). One subject was not administered this subtest due to examiner error. Horizontal lines represent chance performance (solid at 50%) and the 95% confidence limit (dashed at 66%). Data for one child with CIs was unavailable for this task alone.

TABLE 1

Sample Characteristics

Variable	N	Metric	Mean	SD	Range
Child & Family					
Chronologic age	60	Years	10.5	0.8	9.1 – 12.8
WISC Perceptual Reasoning Quotient	60	Standard Score[2]	105.2	12.9	71–135
Highest parent education level	58	# of Years	16.4	1.8	12–20
Family income	59	Category[1]	8.8	1.9	3–10
SES (Education + Income)	58	Composite	25.3	2.6	19–30
CELF – Core Language Index	60	Standard Score[2]	89.19	20.19	42–132
SRSS - Social Skills Rating	60	Standard Score[2]	103.85	14.35	66–129
Audiological Characteristics					
Age at diagnosis	60	Years; Months	0;10	0;8	0;1 – 2;6
Age 1st HA	60	Years; Months	0;11	0;8	0;1 – 2;7
Age at 1st CI	60	Years; Months	1;10	0;7	0;12 – 3;2
Duration 1st CI	60	Years; Months	8;6	0;11	6;2 – 11;2
Age at 2nd CI	29	Years; Months	7;8	1;8	3;10 – 9;11
Pre-Implant Unaided PTA	60	dB HL	108.0	10.9	77–120
Pre-Implant Aided PTA (HA)	60	dB HL	65	15.1	32–80
Current Aided PTA (CI)	60	dB HL	21.5	7.3	4-48
Educational Characteristics					
1st Mainstreamed	60	Grade level[3]	K	1.0	PreK - 6
Class Size >20	60	% of grades	46.5	42.8	0 – 100

[1] Income categories: 1 = Under $5,500; 2 = $ 5,500 – $ 9,999; 3 = $10,000 – $14,999; 4 = $15,000 –$24,999; 5 = $25,000 – $34,999; 6= $35,000 – $49,999; 7 = $50,000 – $64,999; 8 = $65,000 – $79,999; 9 = $80,000 – $94,999; 10= $95,000 – over

[2] Standard Score Mean = 100; SD = 15 (Normal hearing age mates in normative samples) CELF – Clinical Evaluation of Language Fundamentals; SRSS – Social Skills Rating System

[3] Non-parametric summary values reported for grade placement at mainstreaming are the median and inter-quartile range rather than an average and standard deviation.

TABLE 2

Performance on speech perception tests by children with cochlear implants

Linguistic Test	Score	Mean	SD	Range
Lexical Neighborhood Test (LNT)	% correct words at 70 dB SPL	76.7	18.7	0–96
	% correct words at 50 dB SPL	52.0	26.0	0–94
BKB Sentences in Noise	SNR-50 [*] (dB)	9.8	4.6	4–23
Non-word Repetition (CNRep)	% total accuracy	26.9	21.7	0–85
	% correct consonants	73.3	17.4	22–97
	% correct number of syllables	84.7	16.1	20–100
	% correct syllable stress	65.1	19.2	20–100
Indexical Test	**Score**	**Mean**	**SD**	**Range**
Emotion Identification	% correct (chance=25%)	58.5	17.8	25–92
Emotion Discrimination	% correct (chance=50%)	89.8	10.2	54–100
Talker Discrimination Across-Gender	% correct (chance=50%)	89.0	12.6	53–100
Talker Discrimination Within-Female	% correct (chance=50%)	65.4	10.6	41–84

[*] SNR-50 = Signal-to-Noise Ratio at which listener achieves 50%-correct speech understanding

TABLE 3

Inter-correlation matrix of linguistic speech perception measures

	PC	BKB-SIN	NWRcons	LNT-70	LNT-50	NWRsupra
BKB-SIN	-0.917	1	-.809 **	-.829 **	-.746 **	-.679 **
NWRcons	0.920		1	.757 **	.752 **	.763 **
LNT-70	0.895			1	.735 **	.645 **
LNT-50	0.872				1	.644 **
NWRsupra	0.835					1

Explained Variance = 79%

*
p < .01,

**
p < .001

PC: Principal Component Factor Loading

BKB-SIN: Bamford-Kyle-Bench Speech in Noise

NWRcons – non-word repetition: segmental consonant accuracy

LNT – Lexical Neighborhood Test (administered at 50 and at 70 dB SPL)

NWRsupra – non-word repetition: suprasegmental score accuracy of syllable number and stress pattern combined

TABLE 4

Inter-correlation matrix of indexical speech perception measures

	PC	GD	ED	EID	TD
GD	0.825	1	.572 **	.580 **	.415 *
ED	0.839		1	.413 *	.331 *
EID	0.726			1	.608 **
TD	0.745				1

Explained Variance = 62%

*
p < .01,

**
p < .001

PC: Principal Component Factor Loading

GD – Across-Gender Talker Discrimination (male versus female)

ED – Emotion Discrimination

EID – Emotion Identification

TD – Within-Female Talker Discrimination

TABLE 5

Correlations of perception scores with other measured characteristics

	Linguistic Perception	Indexical Perception
Child and Family Characteristics		
Family SES (Ed + Income)	0.120	0.193
Pre Implant Aided PTA	−0.216 *	−0.013
Non-Verbal IQ (WISC[1])	0.030	0.207
Language (CELF)[2]	0.678 ***	0.711 ***
Social Skills (SRSS)[3]	0.186	0.335 **
Cochlear Implant Characteristics		
Age at 1st CI	−0.251 *	−0.372 **
Speech Processor Technology	0.552 ***	0.449 ***
Aided PTA	−0.743 ***	−0.602 ***
Bilateral CI use	0.371 ***	0.362 **
Educational Characteristics		
Grade entered mainstream	−0.374 **	−0.390 **
% Grades class size >20	0.343 **	0.223

*
p<.05;

**
p<.01;

***
p<.001

[1]Wechsler Intelligence Scale for Children;

[2]Clinical Evaluation of Language Fundamentals;

[3]Social Skills Rating System