



Published in final edited form as:

*J Speech Lang Hear Res.* 2009 December ; 52(6): 1390–1400. doi:10.1044/1092-4388(2009/08-0154).

## Spoken Word Recognition in Toddlers Who Use Cochlear Implants

Tina M. Grieco-Calub, Jenny R. Saffran, and Ruth Y. Litovsky

University of Wisconsin–Madison

### Abstract

**Purpose**—The purpose of this study was to assess the time course of spoken word recognition in 2-year-old children who use cochlear implants (CIs) in quiet and in the presence of speech competitors.

**Method**—Children who use CIs and age-matched peers with normal acoustic hearing listened to familiar auditory labels, in quiet or in the presence of speech competitors, while their eye movements to target objects were digitally recorded. Word recognition performance was quantified by measuring each child's *reaction time* (i.e., the latency between the spoken auditory label and the first look at the target object) and *accuracy* (i.e., the amount of time that children looked at target objects within 367 ms to 2,000 ms after the label onset).

**Results**—Children with CIs were less accurate and took longer to fixate target objects than did age-matched children without hearing loss. Both groups of children showed reduced performance in the presence of the speech competitors, although many children continued to recognize labels at above-chance levels.

**Conclusion**—The results suggest that the unique auditory experience of young CI users slows the time course of spoken word recognition abilities. In addition, real-world listening environments may slow language processing in young language learners, regardless of their hearing status.

### Keywords

cochlear implants; spoken word recognition; children

---

Children's proficiency with their native oral language depends to a large extent on their experience during the first few years of life. Early exposure to language provides infants with opportunities to learn about statistical properties (Saffran, Aslin, & Newport, 1996), phonological properties (e.g., Stager & Werker, 1997), word meanings (reviewed in Graf-Estes, Evans, Alibali, & Saffran, 2007; Werker & Yeung, 2005), and grammatical structures (Shi, Werker, & Morgan, 1999). This early experience promotes the development of receptive and expressive language abilities that children continue to fine-tune throughout their life.

Children who are born without hearing have limited exposure to oral language owing to the auditory deprivation imposed by their hearing loss. As a result, these children typically experience delays in the development of spoken language (Yoshinaga-Itano, Sedey, Coulter, & Mehler, 1998). Over the past decade, an increasingly popular intervention strategy has been

---

Contact author: Tina M. Grieco-Calub, University of Wisconsin, Waisman Center, 1500 Highland Avenue, Room 565, Madison, WI 53705. tgriecocalub@niu.edu.

### Disclosure Statement

Ruth Y. Litovsky has consulted and provided written materials for distribution for Cochlear Americas.

to provide infants and young children who are born deaf, or who acquire deafness prior to language development, with cochlear implants (CIs). CIs are surgically implantable devices that provide deaf individuals access to auditory information by converting sound into an electrical signal that is delivered directly to the auditory nerve. Because of technical limitations, the signal delivered to the auditory nerve is spectrally degraded relative to the acoustic input. Despite this issue, advancing CI technology facilitates the development of oral language in deaf individuals. On average, better outcomes are associated with younger ages of implantation (e.g., Kirk et al., 2002; Wang et al., 2008) as well as longer durations of implant experience and better residual hearing prior to implantation (e.g., Nicholas & Geers, 2006). Large individual variability in performance, however, does exist (e.g., Sarant, Blamey, Dowell, Clark, & Gibson, 2001).

Despite the widespread use of CIs, very little is known about how young CI users process spoken language in real time. Because children who use CIs have compromised auditory input, including periods of complete auditory deprivation, the way in which they acquire and process language cannot be assumed to follow the same developmental trajectory as their peers who have normal hearing. In addition, even if CIs are provided by a young age, these devices cannot necessarily ensure “normal” language development because the quality of the auditory input is degraded relative to normal acoustic hearing.

Studies by Houston and colleagues have begun to assess early auditory and language processing in young CI users by extending experimental techniques used with typically developing infants to infants who use CIs. For example, they have used the visual-habituation procedure (Houston, Pisoni, Kirk, Ying, & Miyamoto, 2003) and the preferential looking paradigm (Houston, Ying, Pisoni, & Kirk, 2003) to assess early speech perception in infants who received CIs between 5.8 and 25 months of age. In one of these studies (Houston, Pisoni, et al., 2003), they showed that infants with 1–6 months of CI experience were able to discriminate a repetitive speech sound (“hop hop hop”) from a continuous vowel (“ahh”), suggesting that children can make fundamental distinctions in speech sounds with little auditory experience. Relative to infants with normal hearing, however, these children showed a weaker overall preference for trials containing sound as opposed to silent trials. As discussed by the authors, this latter finding may have significant implications for how young CI users learn and process oral language given that the nuances of speech often provide important cues (e.g., intonation, distributional, and phonotactic) used in word segmentation and word learning.

Speech processing can be further degraded when there is background acoustic competition, commonly referred to as the “cocktail-party effect” (Cherry, 1953). In typically developing infants, the ability to attend to one voice in the presence of competing speech is present to some extent at 5 months of age (Newman, 2005) and is fairly mature by the age of 3 years (Garadat & Litovsky, 2007; Litovsky, 2005). In contrast, individuals with hearing loss (e.g., Dubno, Dirks, & Morgan, 1984; Plomp & Mimpen, 1979) as well as adults and older children who use CIs (Dorman, Loizou, & Fitzke, 1998; Litovsky, Johnstone, & Godar, 2006; Stickney, Zeng, Litovsky, & Assmann, 2004) have difficulty understanding speech when there is background acoustic interference. To date, little is known about the impact of competing acoustic information on speech recognition abilities in very young implant users.

Development is shaped by experience. Thus, it is reasonable to hypothesize that the unique auditory history of children who use CIs influences oral language acquisition. The current study was motivated by the need to understand the mechanisms that are involved in the acquisition and processing of oral language in young CI users. Recently, Marchman and Fernald (2008) showed that the time course of spoken language processing within the second year of life predicts language proficiency in later childhood. There is tremendous individual variability in language outcomes of older children who use CIs, and understanding the ways in which young

CI users process auditory information might reveal possible sources of this wide-ranging variability in later language skills. In addition, identifying mechanisms underlying language acquisition and processing in CI users might provide insight into the neural plasticity of language learning. Beyond its inherent scientific interest, there is clear clinical relevance as caregivers attempt to make informed decisions about intervention strategies for deaf infants.

In an attempt to capture the efficiency of spoken language processing in young CI users, we implemented the looking-while-listening procedure (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998). This is a powerful empirical method that has been used to elucidate speech processing in typically developing toddlers. In the present study, we used this procedure to examine spoken word recognition abilities in two groups of children: 2-year-old children who use CIs and chronologically age-matched peers who have normal hearing. Children were tested using a two-alternative forced-choice task designed to determine whether they could visually recognize known objects after hearing their auditory labels, both in quiet and in the presence of speech competitors (i.e., two-talker babble). Children's eye movements in response to spoken words were coded off-line and analyzed for reaction time and accuracy. Because children with CIs have less experience with spoken English, we hypothesized that their performance would be worse than that of their peers with normal hearing. We also expected that the presence of the speech competitors would reduce word recognition abilities for both groups of children in that similar results have been observed in older children and adults, regardless of hearing status.

## Method

### Participants

Participants consisted of two groups of 2-year-old children. In the CI group ( $n = 26$ ;  $M = 31.3$  months of age), children had severe-to-profound bilateral sensorineural hearing loss and received CIs prior to 29 months of age. In this group, 17 children used bilateral CIs (i.e., one for each ear) and 9 participants used a unilateral CI (i.e., auditory input into one ear). Although all children were primarily using auditory-verbal communication, 14 had manual (i.e., sign) vocabularies. The children were recruited from across the United States and traveled to Madison, Wisconsin, to participate in the experiments (for additional demographic information, see Table 1). Three children were excluded from the analyses because of failure to attend to the task in more than 50% of the trials in each listening condition.

In the normal hearing (NH) group ( $n = 20$ ;  $M = 29.6$  months of age), children had no history of hearing loss, middle ear problems, or other developmental delays per parental report. Three children in this group were excluded from the analyses as a result of fussiness during testing.

### Visual Stimuli

Visual stimuli were digitized photographs of objects that corresponded to the target auditory stimuli. All images were of similar size and placed in  $15'' \times 15''$  white squares that were horizontally aligned on a large screen. The four objects used in this study are among the first words typically learned by English-speaking infants: dog, baby, ball, and shoe. These stimuli have also been used in similar tasks with younger children (e.g., Fernald et al., 1998; Swingley & Aslin, 2000). Two pictures were used for each target object to maintain interest. Objects were also yoked based on their animacy, resulting in two pairs of objects (dog-baby and ball-shoe). Each object served as a target and distractor on an equal number of trials.

### Auditory Stimuli

Target speech stimuli were digitally recorded by a native English-speaking female in an "infant-directed" register. Overall intensity was equalized with customized software created

in MATLAB to yield a level of 65 dB SPL (A-weighted scale). The target stimuli consisted of three frames (*Look at the* [956 ms], *Where is the* [917 ms], and *Can you find it* [702 ms]) and four auditory labels (*doggie* [736 ms], *baby* [750 ms], *ball* [785 ms], and *shoe* [783 ms]).

In addition to target speech stimuli, two-talker babble was created for use in the competitor conditions. Competitor stimuli consisted of sentences from the Harvard IEEE corpus (Rothausser et al., 1969), recorded by a native English-speaking male and scaled to yield an intensity level of 55 dB SPL (A-weighted scale). To create two-talker babble, pairs of sentences were randomly chosen from the list of recorded sentences in the corpus and digitally mixed. When the competitors were combined with the target speech stimuli, the intensity of the target speech was 10 dB greater than the competitors (i.e., +10 dB signal-to-noise ratio). This is a relatively high signal-to-noise ratio relative to those observed in more realistic, noisy environments such as classrooms (e.g., Markides, 1986). We chose to use this more conservative signal-to-noise ratio, however, for two reasons. First, we were unsure as to how young CI users might function in complex auditory environments. Second, this level has been used successfully in tasks with typically developing infants (Newman, 2005; Newman & Jusczyk, 1996).

## Apparatus

The experiment was conducted in an IAC double-walled, sound-attenuated booth. The front and side walls of the booth were covered with drapes that obscured equipment and reduced sound reverberation. The front wall contained an aperture for a digital video camera, two central speakers positioned behind the drapes, and a 58" LCD screen. Visual information was displayed on the screen by an LCD projector. The software program Habit 2000 (Cohen, Atkinson, & Chaput, 2004) was used to present auditory and visual stimuli. The sessions were captured by a DVD recorder (30 frames per second) and coded offline.

## Procedure

The experiment consisted of eight test trials in which each of the four target objects served as the visual target twice, counterbalanced for side of presentation. The experiment was run under two listening conditions: in quiet and in the presence of speech competitors (i.e., competitor). Children in the NH group were randomly assigned to participate in one condition (10 children in the quiet condition and 10 children in the competitor condition). Because children in the CI group spent a number of hours in the lab over the course of 2 days, those children participated in both conditions of the task on 2 different days, with the order of test completion counterbalanced.

Children were seated in a caregiver's lap or in an infant seat that was centered 1 m away from the LCD screen. Caregivers listened to music presented through earphones to eliminate the possibility that they might bias the child's behavior during the experiment. Each test trial began with 3 s of either silence (quiet) or two-talker babble (competitor) prior to the start of the carrier phrase to expose children to the visual stimuli. After every four test trials, children saw a small 8-s movie clip to maintain their interest in the experiment. These trials were not included in the data analysis.

## Coding Eye Movements

The digitally stored video was coded by a trained observer using customized software developed in Anne Fernald's lab at Stanford University. During coding, the observer was "blind" to trial type and side of target picture and was required to judge whether the children were looking at the left object, the right object, or away from either object or were shifting the direction of eye gaze. A subset of the data was coded twice to assess intercoder reliability; 96.8 % of the shifts were within one frame of each other across coders.

## Parental Report of Vocabulary

Parents of the children who use CIs completed the MacArthur Communicative Development Inventory (MCDI), Words and Gestures form (Fenson et al., 1993). Although this version is intended for typically developing children under 16 months of age, it has been shown to be valid in older children who use CIs (Thal, Desjardin, & Eisenberg, 2007).

## Results

The two objectives of this study were to measure the time course of spoken word recognition abilities in young children who use CIs and to assess the effects of a two-talker speech competitor on word recognition performance. The looking-while-listening task (Fernald et al., 1998) was used to examine the time course of word recognition, including accuracy and reaction time (RT; the latency between the label onset and the first eye movement toward the target object), in young children with normal hearing (NH group) and young children who use CIs (CI group).

The time course of spoken word recognition, as illustrated by frame-by-frame data, provides a fine-grained measure of spoken word recognition (see Figure 1). Consistent with prior studies (e.g., Fernald et al., 1998; Swingley & Aslin, 2000), the average proportion of looks at the target object was not different from chance performance during the first 367 ms after the target label onset; children looked equally at the target and distractor objects. After this time point, when eye movements could reflect recognition of the spoken label, the proportion of looking at the target object gradually increased for both groups of children, in quiet as well as in the presence of the speech competitors. Visual inspection of Figure 1 shows that, in quiet, the average shift in eye gaze toward the target object occurred sooner for the NH group than for the CI group. In addition, it appears that the addition of the speech competitors to the task reduced overall performance in both groups of children.

## Accuracy

To compare the accuracy of word recognition across groups and conditions, we computed the proportion of time that each child spent looking at the target object within 367–2,000 ms after the target word onset (Swingley & Aslin, 2000). As shown in Figure 2A, children in the NH group were highly accurate in the quiet condition ( $M \pm SD = 85\% \pm 10\%$ ) but were less accurate when the speech competitors were present ( $73\% \pm 11\%$ ),  $t(18) = 2.5$ ,  $p = .021$ . Children in the CI group showed a similar pattern of overall performance: Accuracy was greater in the quiet condition ( $70\% \pm 13\%$ ) than in the competitor condition ( $58\% \pm 18\%$ ),  $t(25) = 3.1$ ,  $p = .005$ . Although the pattern of performance across the two populations of children was similar, accuracy was poorer in the CI group than in the NH group. The CI group was less accurate both in quiet,  $t(34) = 3.3$ ,  $p = .002$ , and in the presence of the speech competitors,  $t(34) = 2.4$ ,  $p = .02$ .

Children in the CI group completed the looking-while-listening task twice over 2 days, once in quiet and once in the presence of speech competitors (counterbalanced). However, we included two different NH groups; one group was tested in quiet, and the other group was tested in the competitor condition. It is possible that this difference between the protocols used for the two groups of children was responsible for the group differences that were observed. For example, the children in the CI group could have performed more poorly on the task that they completed on the second day of testing (Day 2) because of fatigue or lack of interest. Alternatively, familiarity with the protocol could have led to improved performance on Day 2. To address both of these possibilities, we reran the analyses using only the data from the first testing day (Day 1) for the CI group. Doing so made the results from the two populations of children directly analogous, with a between-subjects manipulation in both cases. The



average accuracy of the children in the CI group who completed the task in quiet on Day 1 ( $70\% \pm 14\%$ ;  $n = 16$ ) continued to be statistically better than the accuracy of their peers who were tested on Day 1 in the presence of the speech competitors ( $54\% \pm 12\%$ ;  $n = 10$ ),  $t(24) = 2.9$ ,  $p = .008$ . This analysis replicates the within-subject analysis and suggests that the competing speech interfered with spoken word recognition in the CI group.

In a subsequent analysis, we compared the Day 1 performance of the CI group with the data from the NH group using a between-subjects analysis of variance. Group (NH or CI) and listening condition (quiet or competitor) served as the two factors. Consistent with the  $t$  tests just described, there were significant main effects of listening condition,  $F(1, 42) = 20.3$ ,  $p < .001$ , and group,  $F(1, 42) = 13.9$ ,  $p = .001$ . The interaction was not statistically significant,  $F(1, 42) = 0.23$ ,  $p = .63$ . Post hoc analyses confirmed that children in the CI group were significantly less accurate than their peers in the NH group both in quiet,  $t(24) = 2.89$ ,  $p = .008$ , and in the presence of the speech competitors,  $t(18) = 3.56$ ,  $p = .002$ . Inspection of individual performance within the CI group (see Figures 2A and 2B), however, reveals a large range of spoken word recognition proficiency: Some children performed similarly to their NH peers, and others performed at chance levels. It is noteworthy, though, that the CI group, on average, recognized spoken words significantly above chance in the quiet condition,  $t(25) = 7.91$ ,  $p < .001$ , consistent with the idea that they were quite capable of performing this task under optimal listening conditions. The inability of the CI group to perform at chance levels in the presence of the speech competitors,  $t(9) = 1.1$ ,  $p = .3$ , however, suggests that the competing acoustic information, even when presented at a conservative signal-to-noise ratio, was detrimental to the speech understanding of these children.

## Reaction Time

The time course of spoken word recognition can also be quantified via RT, or the amount of time children take to visually identify the target object after hearing the target label. RT was defined as the time between the word onset and the first eye movement to the target object within the 367–2,000-ms window defined earlier. Trials were eliminated if eye shifts to the target object occurred within the first 367 ms, or 2,000 ms after the label onset, because those are considered to be uncorrelated with the auditory label (Fernald, Swingley, & Pinto, 2001; Swingley & Aslin, 2000). Trials in which children were fixating on the target object before the onset of the labeling event were also eliminated. In addition, participants who contributed fewer than two trials to their individual average were excluded from the analysis. Based on these criteria, 3 children in the NH group (1 child in the quiet condition and 2 children in the competitor condition) and 15 children in the CI group (6 children in the quiet condition and 9 children in the competitor condition) were removed from the analysis. An additional 3 children in the CI group were excluded because they did not generate an RT in both listening conditions (resulting in a sample size of 14 for this analysis).

RTs for each condition are shown in Figure 3A. In the NH group, average RTs were significantly faster in quiet ( $598 \pm 76$  ms) than in the presence of the speech competitors ( $872 \pm 271$  ms),  $t(15) = 2.9$ ,  $p = .01$ . This suggests that the competing speech slowed down the speech processing skills in this group of typically developing children. Although the CI group showed a similar pattern of performance, with RTs faster in quiet ( $799 \pm 182$  ms) than in the competitor condition ( $943 \pm 259$  ms), the difference between the conditions was not statistically significant,  $t(13) = 1.63$ ,  $p = .127$ . Average RTs in the CI group were slower than those in the NH group in quiet,  $t(21) = 3.1$ ,  $p = .005$ , but not in the presence of the speech competitors,  $t(20) = 0.61$ ,  $p = .55$ .

To control for day of testing, a between-groups analysis including only the data from Day 1 was completed for the CI group (see Figure 3B). This analysis showed that children in the CI group who performed the task in quiet ( $858 \pm 271$  ms;  $n = 13$ ) had significantly faster RTs than

their peers who performed the task in the presence of speech competitors ( $1,097 \pm 266$  ms;  $n = 6$ ),  $t(17) = 2.3$ ,  $p = .035$ . In addition, we compared the Day 1 performance of the CI group with the data from the NH group using a between-subjects analysis of variance. Group (NH or CI) and listening condition (quiet or competitor) served as the two factors. Consistent with the analyses of the accuracy data, there were significant main effects of listening condition,  $F(1, 32) = 13.4$ ,  $p < .001$ , and group,  $F(1, 32) = 11.9$ ,  $p = .002$ . The interaction was not statistically significant,  $F(1, 32) = 0.06$ ,  $p = .8$ . Post hoc analyses replicated the within-subject analysis. Children in the CI group were significantly slower in quiet than their peers in the NH group,  $t(20) = 4.0$ ,  $p = .001$ . When speech competitors were added to the task, however, RTs of the CI group were not statistically different from those of the NH group,  $t(12) = 1.5$ ,  $p = .15$ .

Previous research suggests that the efficiency of speech processing develops with age (Fernald, Perfors, & Marchman, 2006; Fernald et al., 1998). Thus, as children's cognitive abilities emerge and as they have greater experience with spoken language, their ability to process continuous streams of speech improves. Based on this observation, a possible explanation for why the NH group had faster RTs than the CI group is that they simply have more experience with spoken language. Children in the CI group did not receive their implants until 9 to 29 months of age, resulting in a range of auditory experience, or "hearing age," of 6 to 25 months. To control for auditory experience, RTs from the CI group of this study were plotted against published data from acoustic hearing children 15–26 months of age (see Figure 3C). All children in the CI group with a minimum hearing age of 12 months and who had RT data in the quiet condition were included in this comparison ( $n = 18$ ).

As illustrated in Figure 3C, there is a clear improvement in RT between 15 and 26 months of age among the typically developing children assessed in four published studies (Fernald et al., 1998, 2001, 2006; Swingley & Aslin, 2002). However, we do not see the same pattern of results in the current study: The amount of auditory experience (or "hearing age") of the CI group is not correlated with RT ( $r = -.09$ ,  $p = .71$ ). Based on this observation, it appears that factors other than auditory experience are contributing to language processing efficiency in young children who use CIs.

One such factor might be the mode of listening among young CI users. Recently, a number of reports have shown significant functional benefits of bilateral CIs, including improvements in sound localization acuity (Grieco-Calub, Litovsky, & Werner, 2008; Litovsky, Johnstone, & Godar, 2006; Litovsky, Johnstone, Godar, et al., 2006) and better understanding of speech in quiet (Zeitler et al., 2008) or in the presence of competitors such as those used in the present study (Litovsky, Johnstone, & Godar, 2006; Peters, Litovsky, Parkinson, & Lake, 2007). This CI cohort consisted of 17 children who use bilateral CIs and 10 children who use unilateral CIs. Comparison of spoken word recognition accuracy, however, did not reveal a statistically significant difference in performance between these subpopulations, either in quiet,  $t(24) = 0.43$ ,  $p = .67$ , or in the presence of the speech competitors,  $t(24) = 0.42$ ,  $p = .67$ .

### Correlations Between Spoken Word Recognition and Other Variables

Finally, a number of reports have sought to correlate spoken word recognition proficiency with standardized measures of children's expressive and receptive vocabulary (Fernald et al., 2001, 2006; Swingley & Aslin, 2000). In the CI group, productive vocabulary, as quantified by the MCDI, did not correlate with RT in the quiet or competitor conditions. There was, however, a significant correlation between productive vocabulary and spoken word recognition accuracy in quiet ( $r = .489$ ,  $p = .01$ ) as well as between productive vocabulary and hearing age (i.e., amount of auditory experience;  $r = .607$ ,  $p = .001$ ).

Proficiency on the looking-while-listening task, as quantified by spoken word recognition accuracy and RT in quiet or in the presence of speech competitors, did not correlate with

variables such as age of implantation (i.e., duration of CI use), hearing age, chronological age, or mode of listening (i.e., bilateral CIs or unilateral CIs). Early exposure to acoustic hearing also did not correlate with spoken word recognition abilities. For example, 4 of the 26 children in the CI group had some experience with acoustic hearing prior to receiving a CI per parental report. That experience, however, did not appear to be predictive of spoken word recognition abilities because this small subgroup of children had RTs that spanned the range of performance of the larger cohort (e.g., 0.55 to 1.15 ms).

## Discussion

The results of this study showed that children who use CIs are less accurate and slower at visually identifying target objects after hearing the respective auditory labels. In addition, the presence of speech competitors, as is common in naturalistic settings, degraded word recognition abilities in both young children who use CIs and those with normal acoustic hearing. A noteworthy observation, however, is that some of the children in the CI group performed similarly to their age-matched NH peers, both in quiet and in the presence of speech competitors, suggesting that at least a subset of these children have age-appropriate spoken language processing abilities.

Inspection of both individual and group data underscores between-groups differences in lexical processing. As seen in Figures 1 and 3, children in the CI group took significantly longer to fixate the target object in quiet than did their peers in the NH group. The magnitude of this temporal delay is likely to negatively affect language processing in children who use CIs given the rapidity of fluent speech. Based on the comparison made in Figure 3B, it does not appear that amount of experience with spoken language is the sole determinant of language processing skills in these toddlers who use CIs. Consistent with this idea, other potential factors have been noted in the literature, including the age of implantation (Kirk et al., 2002; Sharma, Dorman, & Spahr, 2002), amount of residual hearing prior to implantation (Nicholas & Geers, 2006), and mode of communication (oral alone vs. oral and sign language; e.g., Kirk et al., 2002; Miyamoto, Kirk, Svirsky, & Sehgal, 1999; Svirsky, Robbins, Kirk, Pisoni, & Miyamoto, 2000). In typically developing children, cognitive abilities such as working memory may also play a role (Marchman & Fernald, 2008). It is unclear at the present time which variable dominates. Taken together, it is more likely that there are myriad factors that contribute to overall outcomes.

This study also revealed a detrimental effect of the speech competitors on spoken word recognition in both the NH and CI groups, even at the signal-to-noise ratio of +10 dB used here. The observation that the background speech competitors slowed RTs is consistent with the idea that the ability to segregate different streams of auditory information is still developing in these young language learners (Newman, 2005; Newman & Jusczyk, 1996). This is also consistent with previous studies showing poorer performance of older children who use CIs when speech competitors are present (Litovsky, Johnstone, & Godar, 2006; Peters et al., 2007). It is noteworthy, however, that most of the children in the NH group and some of the children in the CI group performed at better than chance levels on the spoken word recognition task in the presence of the speech competitors. This suggests that the young listeners in this study, even those with degraded auditory experiences, are developing the ability to attend to target speech in quasi-realistic listening situations.

This was an initial investigation of the language processing abilities of young children who are deaf and use CIs. As such, there are a number of caveats to be raised. First, our participants were not a representative sample of all young children who use CIs. Many of the parents of these children were willing to travel long distances to participate in this study. Although this dedication and motivation to participate in research cannot be quantified, it is highly likely that



the parents bring this same level of motivation when implementing the intervention strategies designed to promote spoken language development in their child. As a result, this population of young CI users likely oversamples children with strong spoken language skills. If this argument is valid, the wide range of performance of these children is even more significant and suggests that there are multiple factors that contribute to overall performance. Some of these factors may include variations in audiological and surgical approaches, type of aural rehabilitation, and the extent to which there is exposure to rich sign language in addition to spoken language.

A second issue is that the CI group consisted of children using either a unilateral CI (i.e., auditory input into one ear) or bilateral CIs. Recently, a number of reports have shown significant functional benefits of bilateral CIs, including improvements in sound localization precision in 2-year-olds (Grieco-Calub et al., 2008) and older children (Litovsky, Johnstone, & Godar, 2006; Litovsky, Johnstone, Godar, et al., 2006), and better understanding of speech in the presence of competitors such as those used in the present study (Litovsky, Johnstone, & Godar, 2006; Peters et al., 2007). In this study, we did not observe any significant improvement in word recognition ability in the subset of children who use bilateral CIs, although we did not design the study to directly assess bilateral advantages. As increasingly more infants receive bilateral CIs at younger ages, it will be important to identify potential benefits for language acquisition and language processing in these children. In future experiments, the addition of spatial cues (i.e., changing the location of the background speech) will allow us to better evaluate the benefit of bilateral CIs for this task.

## Summary

Children who use CIs appear to be slower at processing language on this auditory–visual task than their chronologically age-matched peers with normal hearing. The source of this between-groups difference is unclear but may be related to one or more of the variables discussed earlier. In addition, the presence of background speech competitors influences the speed with which young language learners process spoken words. With more language experience, we predict that this influence would diminish. The extent to which this might occur in the CI group, however, is not easily predicted because older CI users also have difficulty understanding speech in adverse listening conditions (e.g., Dorman et al., 1998; Stickney et al., 2004).

This study is one of only a few reports (Houston, Pisoni, et al., 2003; Houston, Ying, et al., 2003) to use objective experimental methodologies to study language processing in very young CI users. Increased use of experimental techniques that allow measurement of online processing, such as those used in this study, in combination with standardized tools, may allow us to identify the mechanisms that underlie language acquisition in young CI users as well as the source of individual variability that leads to the wide range of outcomes in this unique population of children. In addition, objective measures such as the looking-while-listening task may be useful in speech-language pathology clinics to assess speech recognition skills in young children who may be at risk for later language delays. Future experiments similar to the present study are needed to assess the feasibility of implementing such objective measures that have the potential of revealing functional abilities in very young children.

## Acknowledgments

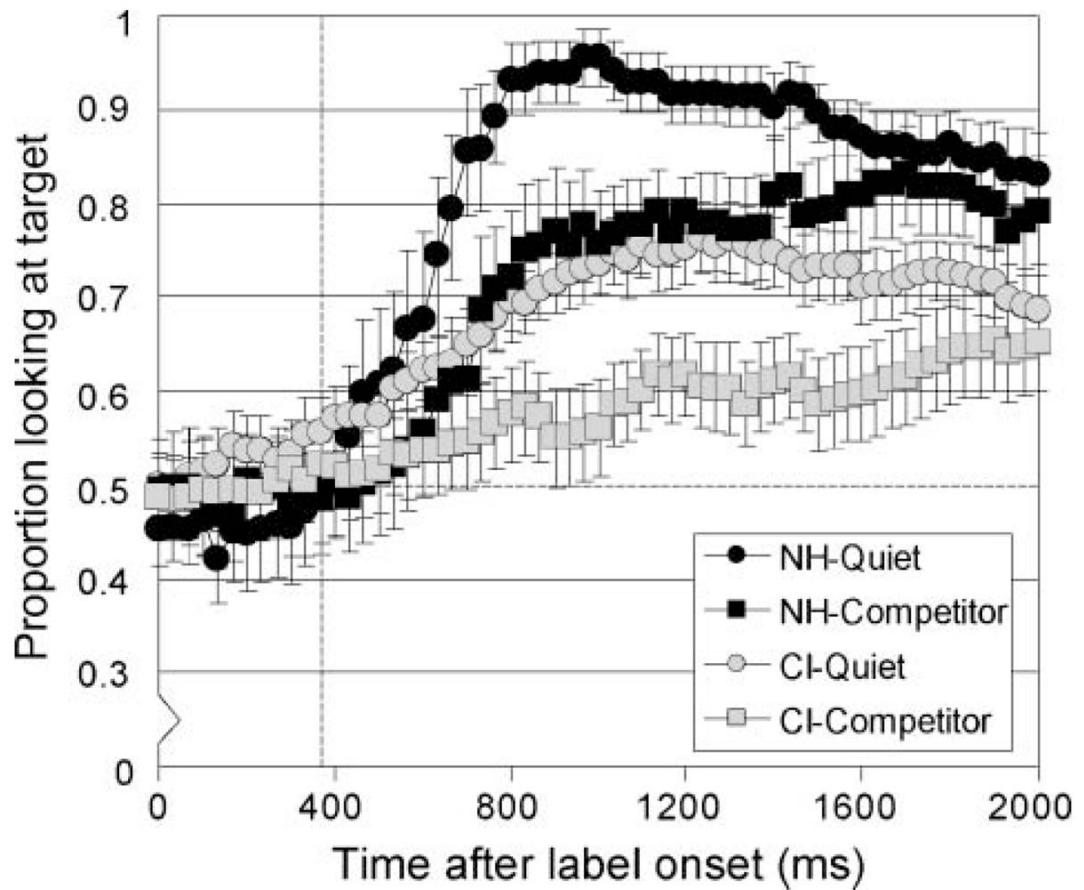
This work was funded by the National Institutes of Health Grant F32DC008452 (awarded to Tina M. Grieco-Calub), Grant R01HD37466 (awarded to Jenny R. Saffran), and Grants R21DC006642 and R01DC008365 (awarded to Ruth Y. Litovsky); National Institute of Child Health and Human Development Core Grant P30HD03352 (awarded to the Waisman Center); and Cochlear Americas, Advanced Bionics, and the Med-El Corporation (sponsoring family travel and participation). One of the authors (Ruth Y. Litovsky) has consulted and provided written materials for distribution for Cochlear Americas.

We would like to thank the following people for their contributions: Katherine Graf Estes for help in experimental design; Shelly Godar, Tanya Jensen, Eileen Storm, Susan Richmond, Molly Bergsbaken, and Kristine Henslin for help in data collection; and Derek Houston, Karen Kirk, and Alexa Romberg for helpful comments on the article.

## References

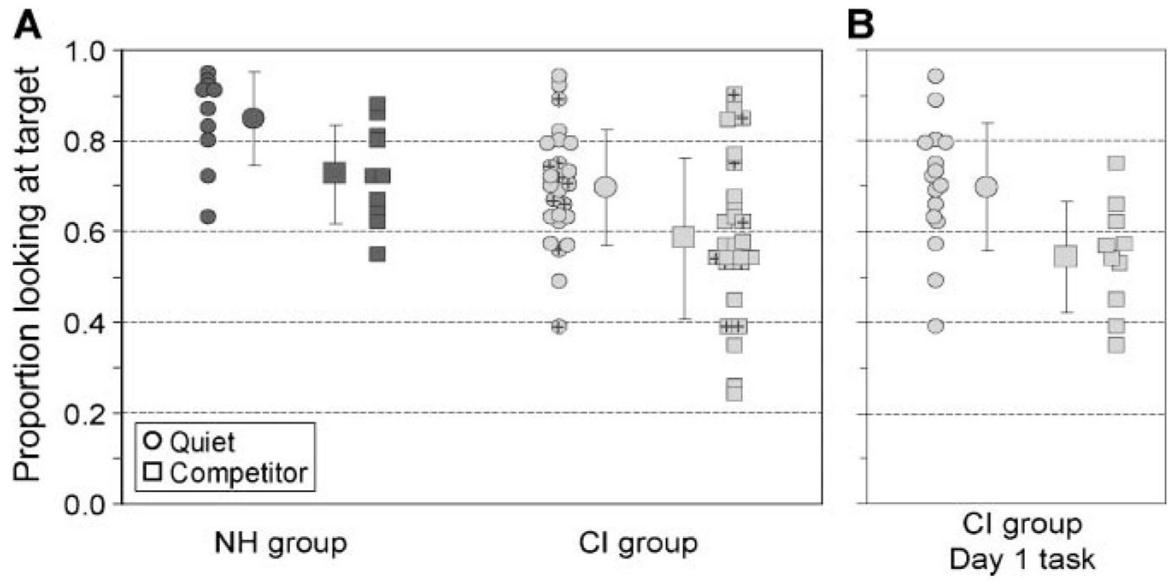
- Cherry EC. Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America* 1953;25:975–979.
- Cohen, LB.; Atkinson, DJ.; Chaput, HH. *Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0)*. Austin: University of Texas; 2004.
- Dorman MF, Loizou PC, Fitzke J. The identification of speech in noise by cochlear implant patients and normal-hearing listeners using 6-channel signal processors. *Ear and Hearing* 1998;19:481–484. [PubMed: 9867296]
- Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America* 1984;76:87–96. [PubMed: 6747116]
- Fenson, L.; Dale, PS.; Reznick, JS.; Thal, D.; Bates, E.; Hartung, JP., et al. *The MacArthur-Bates Communicative Development Inventories: User's guide and technical manual*. Baltimore: Brookes; 1993.
- Fernald A, Perfors A, Marchman VA. Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology* 2006;42:98–116. [PubMed: 16420121]
- Fernald A, Pinto JP, Swingley D, Weinberg A, McRoberts GW. Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science* 1998;9:228–231.
- Fernald A, Swingley D, Pinto JP. When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development* 2001;72:1003–1015. [PubMed: 11480931]
- Garadat SN, Litovsky RY. Speech intelligibility in free field: Spatial unmasking in preschool children. *Journal of the Acoustical Society of America* 2007;121:1047–1055. [PubMed: 17348527]
- Graf Estes K, Evans JL, Alibali MW, Saffran JR. Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science* 2007;18:254–260. [PubMed: 17444923]
- Grieco-Calub TM, Litovsky RY, Werner LA. Using the observer-based psychophysical procedure to assess localization acuity in toddlers who use bilateral cochlear implants. *Otology & Neurotology* 2008;29:235–239. [PubMed: 18180690]
- Houston DM, Pisoni DB, Kirk KI, Ying EA, Miyamoto RT. Speech perception skills of deaf infants following cochlear implantation: A first report. *International Journal of Pediatric Otorhinolaryngology* 2003;67:479–495. [PubMed: 12697350]
- Houston DM, Ying EA, Pisoni DB, Kirk KI. Development of pre- word-learning skills in infants with cochlear implants. *Volta Review* 2003;103:303–326.
- Kirk KI, Miyamoto RT, Lento CL, Ying E, O'Neill T, Fears B. Effects of age at implantation in young children. *Annals of Otology, Rhinology & Laryngology* 2002;119(Suppl):69–73.
- Litovsky RY. Speech intelligibility and spatial release from masking in young children. *Journal of the Acoustical Society of America* 2005;117:3091–3099. [PubMed: 15957777]
- Litovsky RY, Johnstone PM, Godar SP. Benefits of bilateral cochlear implants and/or hearing aids in children. *International Journal of Audiology* 2006;45(Suppl 1):S78–S91. [PubMed: 16938779]
- Litovsky RY, Johnstone PM, Godar S, Agrawal S, Parkinson A, Peters R, et al. Bilateral cochlear implants in children: Localization acuity measured with minimum audible angle. *Ear and Hearing* 2006;27:43–59. [PubMed: 16446564]
- Marchman VA, Fernald A. Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science* 2008;11:F9–F16. [PubMed: 18466367]
- Markides A. Speech levels and speech-to-noise ratios. *British Journal of Audiology* 1986;20:115–120. [PubMed: 3719159]
- Miyamoto RT, Kirk KI, Svirsky MA, Sehgal ST. Communication skills in pediatric cochlear implant recipients. *Acta Oto-Laryngologica* 1999;119:219–224. [PubMed: 10320080]

- Newman RS. The cocktail party effect in infants revisited: Listening to one's name in noise. *Developmental Psychology* 2005;41:352–362. [PubMed: 15769191]
- Newman RS, Jusczyk PW. The cocktail party effect in infants. *Perception & Psychophysics* 1996;58:1145–1156. [PubMed: 8961826]
- Nicholas JG, Geers AE. Effects of early auditory experience on the spoken language of deaf children at 3 years of age. *Ear and Hearing* 2006;27:286–298. [PubMed: 16672797]
- Peters BR, Litovsky R, Parkinson A, Lake J. Importance of age and postimplantation experience on speech perception measures in children with sequential bilateral cochlear implants. *Otology & Neurotology* 2007;28:649–657. [PubMed: 17712290]
- Plomp R, Mimpen AM. Speech-reception threshold for sentences as a function of age and noise level. *Journal of the Acoustical Society of America* 1979;66:1333–1342. [PubMed: 500971]
- Rothausser E, Chapman W, Guttman N, Nordby K, Silbiger H, Urbanek G, et al. IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio Electroacoustics* 1969;17:225–246.
- Saffran JR, Aslin RN, Newport EL. Statistical learning by 8-month-old infants. *Science* 1996 December 13;274:1926–1928. [PubMed: 8943209]
- Sarant JZ, Blamey PJ, Dowell RC, Clark GM, Gibson WP. Variation in speech perception scores among children with cochlear implants. *Ear and Hearing* 2001;22:18–28. [PubMed: 11271973]
- Sharma A, Dorman MF, Spahr AJ. A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear and Hearing* 2002;23:532–539. [PubMed: 12476090]
- Shi R, Werker JF, Morgan JL. Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition* 1999;72:B11–B21. [PubMed: 10553673]
- Stager CL, Werker JF. Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature* 1997 July 24;388:381–382. [PubMed: 9237755]
- Stickney GS, Zeng FG, Litovsky R, Assmann P. Cochlear implant speech recognition with speech maskers. *Journal of the Acoustical Society of America* 2004;116:1081–1091. [PubMed: 15376674]
- Svirsky MA, Robbins AM, Kirk KI, Pisoni DB, Miyamoto RT. Language development in profoundly deaf children with cochlear implants. *Psychological Science* 2000;11:153–158. [PubMed: 11273423]
- Swingle D, Aslin RN. Spoken word recognition and lexical representation in very young children. *Cognition* 2000;76:147–166. [PubMed: 10856741]
- Swingle D, Aslin RN. Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science* 2002;13:480–484. [PubMed: 12219818]
- Thal D, Desjardin JL, Eisenberg LS. Validity of the MacArthur-Bates Communicative Development Inventories for measuring language abilities in children with cochlear implants. *American Journal of Speech-Language Pathology* 2007;16:54–64. [PubMed: 17329675]
- Wang NY, Eisenberg LS, Johnson KC, Fink NE, Tobey EA, Quittner AL, et al. Tracking development of speech recognition: Longitudinal data from hierarchical assessments in the Childhood Development After Cochlear Implantation Study. *Otology & Neurotology* 2008;29:240–245. [PubMed: 18223451]
- Werker JF, Yeung HH. Infant speech perception bootstraps word learning. *Trends in Cognitive Sciences* 2005;9:519–527. [PubMed: 16202639]
- Yoshinaga-Itano C, Sedey AL, Coulter DK, Mehl AL. Language of early- and later-identified children with hearing loss. *Pediatrics* 1998;102:1161–1171. [PubMed: 9794949]
- Zeitler DM, Kessler MA, Terushkin V, Roland TJ Jr, Svirsky MA, Lalwani AK, et al. Speech perception benefits of sequential bilateral cochlear implantation in children and adults: A retrospective analysis. *Otology & Neurotology* 2008;29:314–325. [PubMed: 18494140]



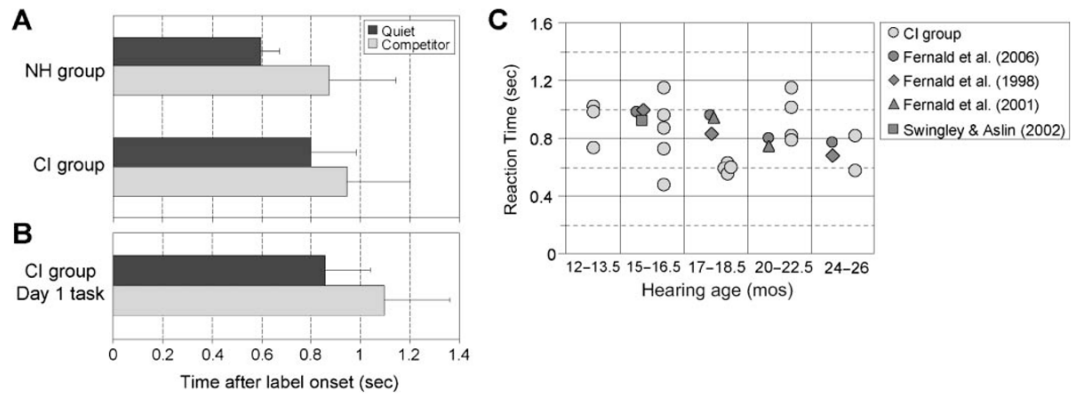
**Figure 1.**

Time course of spoken word recognition. Children with normal hearing (NH; dark gray) and children who use cochlear implants (CI; light gray) performed a looking-while-listening task in quiet (circles) and in the presence of speech competitors (squares). Dotted line represents 367 ms.



**Figure 2.**  
A: Accuracy of spoken word recognition. Accuracy is defined as the proportion of frames in which children are looking at the target object within 367–2,000 ms after the label onset. Individual and mean ( $\pm SD$ ) data are plotted for the normal hearing (NH) group (dark gray), children in the cochlear implant (CI) group who use bilateral devices (empty light gray), and children in the CI group who use a unilateral device (“+” light gray). Scores closest to a proportion of 1 represent longer looks at the target object (i.e., high accuracy), and scores closest to a proportion of 0 represent longer looks at the distractor object (i.e., low accuracy).  
B: Individual and mean ( $\pm SD$ ) accuracy of the CI group for tasks completed on Day 1 of testing.





**Figure 3.**

A: Reaction time (RT) of spoken word recognition. RT is defined as the time between the label onset and the first eye movement toward the target object ( $M \pm SD$ ). B: Average RTs of the CI group for tasks completed on Day 1 of testing. C: Effect of auditory experience. RT data from individual CI group children were plotted with average RT data from four published studies of children with normal hearing matched for auditory experience.

Table 1

Participant demographics.

Participant	Gender	Age at visit (mo)	Age at 1st CI (mo)	Experience with 1st CI at visit (mo)	Age at 2nd CI (mo)	Experience with 2nd CI at visit (mo)	Etiology	1st CI device (ear)	2nd CI device (ear)	Speech coding strategy
<u>Bilateral CIs</u>										
CIBV	M	29	17	12	23.5	6.0	Connexin 26, 30	HiRes (right)	HiRes (left)	HiRes-S
CIBX	M	31.5	10	21.5	12	19.5	Waardensburg	Nucleus 24 Contour (right)	Nucleus 24 contour (left)	ACE
CICA	M	35	29	6	29	6.0	Unknown	Pulsar (simultaneous)	Pulsar (simultaneous)	CIS+
CICB	F	31	10.5	20.5	25	6.0	Connexin 26	Nucleus 24 Contour (right)	Freedom Contour (left)	ACE
CICF	F	34.5	18	16.5	28.5	6.0	Bacterial meningitis at age 1 mo	Freedom Contour (right)	Freedom Contour (left)	ACE
CICH	M	26	12	14	12	15	Unknown	Pulsar (simultaneous)	Pulsar (simultaneous)	CIS+
CICK	M	28.5	13	15.5	15.5	13	Connexin 26	HiRes (right)	HiRes (left)	HiRes-P
CICP	F	26.5	13	13.5	20	6.5	Connexin 26	Freedom Contour (right)	Freedom Contour (left)	ACE
CICR	M	25.5	12.5	13	16.5	9.5	Connexin 26	HiRes (left)	HiRes (right)	HiRes-P
CICS	M	34.5	18	16.5	18	16.5	Connexin 26	HiRes (simultaneous)	HiRes (simultaneous)	HiRes-P
CIDC	M	31	18	13	18	13	Unknown	HiRes (simultaneous)	HiRes (simultaneous)	HiRes-S
CIDD	F	36	21	15	29	7	Unknown	Freedom Contour (left)	Freedom Contour (right)	ACE
CIDE	M	26.5	13	13.5	13	13.5	Unknown	HiRes (simultaneous)	HiRes (simultaneous)	HiRes-P
CIDH	M	35.5	17	18.5	26	9.5	Unknown	Freedom Contour (right)	Freedom Contour (left)	ACE
CIDK	M	26.5	10.5	16	10.5	16	Connexin 26	Freedom Contour (simultaneous)	Freedom Contour (simultaneous)	ACE
CIDL	F	27.5	14	13.5	18	9.5	Connexin 26	Pulsar (right)	Sonata (left)	CIS+
CIDM	F	31.5	13	18.5	25	6.5	Unknown	HiRes (left)	HiRes (right)	HiRes-P
<b>M</b>		<b>30.4</b>	<b>15.2</b>	<b>15.2</b>	<b>20.0</b>	<b>10.6</b>				
<u>Unilateral CIs</u>										
CICC	M	36	15	21			Familial deafness	Nucleus 24 Contour (left)		ACE
CICE	M	30	10.5	19.5			Connexin 26, 30	HiRes (right)		HiRes-P
CICI	M	36	14.5	21.5			Unknown	Pulsar (right)		CIS+
CICJ	M	35	10	25			Connexin 26	HiRes (right)		HiRes-P
CICL	M	27	16	11			Connexin 26	Freedom Contour (right)		ACE
CICM	M	36	13	23			Connexin 26	HiRes (right)		HiRes-S

Participant	Gender	Age at visit (mo)	Age at 1st CI (mo)	Experience with 1st CI at visit (mo)	Age at 2nd CI (mo)	Experience with 2nd CI at visit (mo)	Etiology	1st CI device (ear)	2nd CI device (ear)	Speech coding strategy
C1CN	F	32.5	15.5	17			Connexin 26	Freedom Contour (right)		ACE
C1CQ	M	31	14.5	16.5			Usher syndrome, Type I	Freedom Contour (right)		ACE
C1DB	F	35	15.5	19.5			Unknown	HiRes (right)		HiRes-P
<i>M</i>		<b>33.2</b>		<b>19.3</b>						

*Note.* Cochlear implant (CI) devices: Freedom, Nucleus 24 (Cochlear Americas); HiRes (Advanced Bionics); Pulsar, Sonata (Med-El Corporation).