JSLHR

Article

Speech Perception in Noise by Children With Cochlear Implants

Amanda Caldwell[a] and Susan Nittrouer[a]

Purpose: Common wisdom suggests that listening in noise poses disproportionately greater difficulty for listeners with cochlear implants (CIs) than for peers with normal hearing (NH). The purpose of this study was to examine phonological, language, and cognitive skills that might help explain speech-in-noise abilities for children with CIs.
Method: Three groups of kindergartners (NH, hearing aid wearers, and CI users) were tested on speech recognition in quiet and noise and on tasks thought to underlie the abilities that fit into the domains of phonological awareness, general language, and cognitive skills. These last measures were used as predictor variables in regression analyses with speech-in-noise scores as dependent variables.
Results: Compared to children with NH, children with CIs did not perform as well on speech recognition in noise or on most other measures, including recognition in quiet. Two surprising results were that (a) noise effects were consistent across groups and (b) scores on other measures did not explain any group differences in speech recognition.
Conclusions: Limitations of implant processing take their primary toll on recognition in quiet and account for poor speech recognition and language/phonological deficits in children with CIs. Implications are that teachers/clinicians need to teach language/phonology directly and maximize signal-to-noise levels in the classroom.

Key Words: cochlear implants, speech perception, children, speech recognition in noise and quiet

Recognizing speech in noisy conditions has always been viewed as presenting special difficulty for listeners with hearing loss. Before cochlear implants (CIs), when hearing aids (HAs) were the only means of amplification, this difficulty could be explained largely by the fact that abnormal cochlear function is associated with broadened auditory filters (e.g., Glasberg & Moore, 1986; Leek, Dorman, & Summerfield, 1987). Phonemically relevant spectral structure, such as formant frequencies, is poorly represented with broadened filters (e.g., Revoile, Pickett, & Kozma-Spytek, 1991). As a result, perceptual segregation of that spectral structure from background noise is difficult (Baer, Moore, & Gatehouse, 1993; Bernstein & Brungart, 2011; Boothroyd, Mulhearn, Gong, & Ostroff, 1996; Fu, Shannon, & Wang, 1998). The development of CIs as a treatment for severe to profound hearing loss has done nothing to improve this particular problem because processing strategies in these devices are designed to deliver only spectrally broad signals to the auditory system (e.g., Rubinstein, 2004).

CIs have, however, improved speech recognition *in quiet* for listeners with severe to profound hearing loss, with average isolated word recognition scores for adults of 40% to 50% correct (Firszt et al., 2004) and some individual implant users scoring much higher; however, listeners using CIs continue to struggle in noisy situations (e.g., Zeng & Galvin, 1999). In addition to the problems introduced by processing strategies that provide only a broad representation of phonemically relevant spectral structure, a lack of temporal fine structure contributes to the difficulty these listeners experience segregating a target signal from background noise (Lorenzi, Gilbert, Carn, Garnier, & Moore, 2006). From these constraints on the kinds of signal structure available through implants, it is easy to appreciate why listeners with CIs might struggle to recognize speech in noisy environments. There are, however, factors other than the constraints on signal structure affecting bottom-up auditory processes that influence how well listeners recognize speech in noisy backgrounds.

[a]The Ohio State University, Columbus
Correspondence to Amanda Caldwell: amanda.caldwell@osumc.edu

Editor: Sid Bacon
Associate Editor: Emily Tobey
Received December 8, 2011
Accepted June 1, 2012
DOI: 10.1044/1092-4388(2012/11-0338)

Top-Down Linguistic Effects

A listener's knowledge of linguistic structure and the ability to bring that knowledge to bear on speech perception affect how well one recognizes speech in noise. Several metrics have been developed to quantify the contribution made by these top-down influences. In particular, Boothroyd (Boothroyd, 1968; Boothroyd & Nittrouer, 1988) developed such a metric, known as the *j* factor, and it has been used on data for listeners from age 4 years to elderly (Nittrouer & Boothroyd, 1990). When it comes to word recognition, this metric is built on the principle that the probability of recognizing a word is related to the probabilities of recognizing the separate constituents of that word, a relationship that can be represented by this equation:

$$p_w = p_p^n, \tag{1}$$

where p_w is the probability of recognizing the word; p_p is the probability of recognizing each part, or phoneme; and n is the number of phonemes in the word. Of course, a listener's lexical knowledge influences word recognition such that the probability of recognizing a word does not depend on the recognition of each phoneme separately, as would be the case for nonsense items. Therefore, Equation 1 can be changed as follows:

$$p_w = p_p^j, \tag{2}$$

where j is the number of independent channels of information required for recognition and is between 1 and n. Equation 2 can now be rewritten to provide the metric of the effective number of information channels needed to recognize a word:

$$j = \log(p_w)/\log(p_p). \tag{3}$$

According to this formula, j is a dimensionless factor that serves as an index of how strongly lexicality influences word recognition. The smaller j is, the greater the effect of lexicality on recognition. This effect is the combination of having the requisite lexical knowledge and of being able to bring that knowledge to bear on recognition. Nonetheless, early investigations of top-down linguistic constraints on speech perception typically assumed normal language knowledge and instead focused on measuring the extent to which listeners use that knowledge in perception (Hirsh, Reynolds, & Joseph, 1954; Miller, Heise, & Lichten, 1951; Pollack, Rubenstein, & Decker, 1959). Studies with other populations, however, are reminders that the first component of the effect cannot be assumed. Listeners such as second-language learners are perfectly capable of using linguistic constraints during perception, if functioning in their first language, but lack adequate familiarity with linguistic and phonological structure in their second language. As a result, speech perception under difficult listening conditions, such as noisy environments, is hard for these listeners (Bradlow & Pisoni, 1999; Cutler, Garcia Lecumberri, & Cooke, 2008; Flege, MacKay, & Meador, 1999; Pinet & Iverson, 2010; Rogers, Lister, Febo, Besing, & Abrams, 2006; von Hapsburg, Champlin, & Shetty, 2004). When evaluating speech perception in noise for clinical populations, it is possible that a lack of linguistic knowledge might similarly explain some part of any observed deficit that is found for those listeners.

Lessons From Individuals With Dyslexia

One population other than listeners with hearing loss that demonstrates particular difficulty recognizing speech in noise consists of individuals with reading problems. In 1983, Brady, Shankweiler, and Mann showed that children with dyslexia recognized words in noise more poorly than their peers with normal hearing (NH), even though recognition scores in quiet were comparable across groups. This finding was attributed to poor phonological category formation on the part of the children with dyslexia, and work since then has supported that suggestion (e.g., Serniclaes, Ventura, Morais, & Kolinsky, 2005; Vance & Martindale, 2011). Other studies have found that individuals with reading problems have difficulty creating categories from sensory inputs, regardless of whether they are related to speech (Ahissar, Lubin, Putter-Katz, & Banai, 2006; Nittrouer, Shune, & Lowenstein, 2011). In turn, this problem has been identified as a source of the speech-perception-in-noise difficulties exhibited by this group. The ability to make strong predictions about what kinds of structures to expect makes perception resistant to noise masking (Ahissar, 2007). From this perspective, the problems dyslexic children encounter listening to speech in noise could be described as strictly top-down effects: Because of their weakly defined phonological categories, they are ill equipped to predict the structures that would be in the noise. This explanation could similarly explain the finding that second-language learners have greater difficulty recognizing speech in noise than first-language learners: Because second-language listeners have more weakly specified phonological categories as well, they would also be less capable of predicting phonological form through the interfering noise. In any event, it is unlikely that the deficits that listeners with dyslexia exhibit when it comes to listening to speech in noise are due to problems at the auditory periphery. Ziegler and colleagues (Ziegler, Pech-Georgel, George, Alario, & Lorenzi, 2005; Ziegler, Pech-Georgel, George, & Lorenzi, 2011) conducted two studies with a slightly different population (individuals with specific language impairments) and found that these listeners did indeed have more difficulty than typically developing children understanding speech in a variety of noises. However, the children with language impairment showed masking release for fluctuating, rather than stationary, noise comparable in magnitude to that of typically developing children. Because that kind of masking release is a peripheral effect, their enhanced masking per se is unlikely due to auditory deficits at the periphery.

Another deficit often attributed to children with dyslexia is one described as a *temporal processing* deficit (Gaab, Gabrieli, Deutsch, Tallal, & Temple, 2007; Merzenich et al., 1996; Tallal, 1980; Tallal, Miller, & Fitch, 1993; Tallal & Piercy, 1973, 1974). This proposed deficit is commonly viewed as a problem in how the auditory system handles rapidly changing acoustic structure, and the proposed effect is that it hinders the ability of children with dyslexia to develop highly refined phonological categories. Consequently, it could be a source of their problems in understanding speech in noise. One kind of phonetic contrast that has been used to demonstrate temporal processing deficits in children with dyslexia is voice onset time (VOT): Children with reading problems have exhibited shallower labeling functions for stimuli along VOT continua than children without such deficits (Breier et al., 2001; Godfrey, Syrdal-Lasky, Millay, & Knox, 1981; Manis et al., 1997). VOT is a particularly good property to use in investigations of whether children with hearing loss who use CIs have difficulty forming phonemic categories, similar to the problems found for children with dyslexia. This temporal structure should be well represented in implants because processing strategies recover envelopes in a number of channels and code this information onto the electrodes. Thus, this acoustic cue of VOT could offer a test of whether there is an underlying problem with category formation for children with CIs.

Children Who Use CIs

Given all that is known about what it takes to recognize speech in noisy backgrounds, it comes as no surprise to find that children with CIs have been found to have more trouble than other children doing so (Davidson, Geers, Blamey, Tobey, & Brenner, 2011). Peripheral processes that allow listeners with normal cochlear function to separate target signals from background noise are impaired in these children, because of the signal-processing limitations of the devices. Some of the perceptual and cognitive processes that seem to influence speech recognition in noisy environments from a top-down direction are also impaired, likely adding to these listeners' problems understanding in noise. Children with CIs do not have phonemic categories that are as well formed as their peers with NH (Geers & Hayes, 2011; Johnson & Goswami, 2010), a factor that could inhibit their abilities to predict the signal. Although implants are associated with tremendous improvements in spoken language abilities for children with hearing loss over what they are capable of without them (Tomblin, Spencer, Flock, Tyler, & Gantz, 1999), these children continue to show impaired abilities. Compared to their peers with NH, children with CIs continue to struggle at recognizing language structure in general (Niparko et al., 2010; Nittrouer, 2010) and at building vocabularies (Hayes, Geers, Treiman, & Moog, 2009; James, Rajput, Brinton, & Goswami, 2009; Nittrouer, 2010). These broader language deficits would reasonably be expected to affect the abilities of children with CIs to recognize speech in noise. Finally, children with CIs have demonstrated impaired abilities in processes fitting under the heading of cognitive skills, in particular working memory (Burkholder & Pisoni, 2003; Cleary, Pisoni, & Geers, 2001; Pisoni, Kronenberger, Roman, & Geers, 2011). Deficits in cognitive skills have not specifically been implicated in problems recognizing speech in noise. However, it can be assumed that difficulties forming and retaining a memory of what was heard could influence outcomes.

The Current Study

Before this study was conducted, it was anticipated that children with CIs would show poorer speech perception in noise than children with NH, or even children with enough residual hearing to use HAs. The study we report here was undertaken mainly to examine factors that might account for variability among children with CIs in their abilities to recognize speech in noise, thus improving researchers' collective understanding of this deficit. In particular, three measures of phonological awareness were obtained, along with a measure of speech perception in noise. The three measures of phonological awareness spanned a range of developmental aptitude, thus increasing the likelihood that maximum variability among participants would be found. Such variability is necessary for performing powerful regression analyses. Other factors examined as potential predictors of children's abilities to recognize speech in noise included their abilities to form phonemic categories based on a temporal property (VOT). Examining children's labeling of stimuli along a VOT continuum provided an estimate of how well these children can use rapidly changing acoustic structure of a sort likely to be available through an implant. It also permitted insight into how well deaf children with CIs can form categories when given access to the sensory information on which those categories are based. Also evaluated in the current study were children's language comprehension, vocabulary size, short-term memory, and speed of perceptual processing. All participants with hearing loss, and most with NH, in this study were part of a longitudinal investigation and had been tested repeatedly at regular intervals since infancy. This fact made them a particularly good sample to study, because no evidence to suspect secondary handicaps that might interfere with the development of spoken language has been found for any child in the longitudinal study.

Method

Participants

Fifty-four children who had just completed kindergarten came to The Ohio State University during the summer of 2010 to participate in this study. Of these, 35 had permanent sensorineural hearing loss with three-frequency pure-tone averages greater than 50 dB hearing level in the better ear. Twenty-seven of those children had severe to profound sensorineural hearing loss and wore one or two CIs. Eight had moderate hearing loss and wore bilateral HAs. Another 19 children had NH. Pure-tone audiometric measurements made at the time of testing confirmed these category descriptions. All children except for four in the NH group had participated in a longitudinal study from 12 to 48 months of age (Nittrouer, 2010). All children with hearing loss in the study received intervention services focused on spoken language development starting shortly after their hearing loss was identified.

Demographic Measures

Demographic information for the three groups is provided in Table 1. Range of age at the time of testing was 73–85 months for children with NH, 72–83 months for children with HAs, and 73–94 months for children with CIs. One child with CIs had to repeat kindergarten, which explains the greater age range for that group. Socioeconomic status was indexed using a two-factor scale on which both the highest educational level and the occupational status of the primary income earner in the home are considered (Nittrouer & Burton, 2005). Scores for each of these factors range from 1 to 8, with 1 being the lowest and 8 being the highest. Values for the two factors are multiplied together, resulting in a range of possible scores from 1 to 64. In general, a score of 30 represents a household in which the primary income earner has a 4-year university degree and a job such as a mid-level manager or a teacher. Scores of 20 represent households in which the primary income earner has a high school diploma and works in a service industry, in construction, or as a skilled craftsman. The scores obtained for these children suggest they all had reasonably rich language environments in the home.

Table 1. Mean demographic measures.

	Group					
	NH		HA		CI	
Measure	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age at time of testing (months)	78	3	78	4	81	5
Proportion of boys	.42		.50		.48	
Socioeconomic status (score)	36	13	25	11	33	12
Age at identification (months)			9	11	8	8
Pre-implant (CIs)/current (HAs) PTA			65	11	99	18
Age at first implant (months)					21	13
Age at second implant (months; *n* = 18)					35	14
Mean length of first implant use (months)					61	13

Note. Except where noted, numbers in each group are 19 for the group with normal hearing (NH), eight for the group that wore hearing aids (HAs), and 27 for the group with cochlear implants (CIs). PTA = pure-tone average.

The bottom four rows of Table 1 show demographic information for the children with hearing loss. All children with hearing loss had been identified before 2 years of age, and most had been identified before 1 year. The children with CIs had received those implants early, which for 21 of the 27 children meant before age 2. Consequently, these children had considerable experience using their implants. Eighteen of the children with implants wore two. Thirteen children with implants had worn an HA on the opposite ear (i.e., bimodal experience) for a period of a year or more.

Equipment

All testing took place in sound-attenuated rooms. All stimuli used in testing were presented via a computer with a Creative Labs Soundblaster digital-to-analog card using a 44.1-kHz sampling rate with 16-bit digitization and a Roland MA-12C-powered speaker for audio presentation, placed 1 m in front of the child at a 0° azimuth. The phonological awareness tasks were presented in audiovisual format using a 1,500-kbps data rate and 24-bit digitization for video presentation. Presentation for speech recognition in noise was audio only.

All test sessions were video-recorded using a Sony HDR-XR550V video recorder so that scoring could be done later. Children wore Sony FM transmitters in specially designed vests that transmitted speech signals to the receivers. Those receivers provided direct line input to the hard drives of the cameras to ensure good sound quality for all recordings.

General Procedure

Children arrived in Columbus, Ohio, the night before testing started. Four to six children attended each camp. They were tested individually in six separate sessions over a 2-day period. Each test session consisted of several tasks that lasted no longer than 1 hr. Children had a minimum of 1 hr between test sessions. Several kinds of measures were collected to assess children's abilities to recognize words and constituent phonemes in noise and to evaluate the factors that likely affect those recognition abilities. All procedures were approved by the institutional review board of The Ohio State University.

Stimuli and Task-Specific Procedures

Speech in noise. Eighteen word lists from Mackersie, Boothroyd, and Minniear (2001) were used. Each list consisted of 10 phonetically balanced CVC words. Noise with a flat spectrum was generated using a random-noise generator. Six word lists were presented at each of three signal-to-noise ratios (SNRs): –3 dB, 0 dB, and +3 dB. These levels were chosen because they have been used often in earlier work investigating speech recognition in noise by children, both typically developing and with dyslexia (Brady et al., 1983; Nittrouer et al., 2011; Nittrouer & Boothroyd, 1990). Although the same 18 lists were presented to all children, the SNR at which each list was presented varied randomly across children. The order of presentation of lists, and so of SNRs, also varied randomly, with the stipulation that two lists could not be presented at the same SNR consecutively. During presentation, the level of the words was held constant at 68 dB SPL, and the level of noise varied. After testing in noise was completed with these 18 word lists, the same words were presented in quiet for recognition. Dependent measures were the percentages of correct words and phonemes.

The *j* factor (Boothroyd & Nittrouer, 1988) was used to index the contribution of lexical effects to word recognition. The *j* factor is not reliable when percentages of phonemes or words recognized correctly are below 5% or above 95%. In this study, these factors were not available for some children with HAs or CIs in some SNRs because they scored below 5% correct on either phonemes or words. Using the *j* factor permitted the examination of whether any potential group differences observed for recognition scores might be due to differences in the extent to which children made use of lexical effects during speech perception.

In a separate session, the CID W-22 word lists (Hirsh et al., 1954) were presented in quiet. These lists consist of 50 words each and are commonly administered in clinical settings. There are four lists, and the lists presented to individual children varied. These additional materials were used to gauge how ecologically valid the Mackersie et al. (2001) materials are for use with children.

Phonological awareness. Three tasks assessing phonological awareness were used in order to cover a broad range of developmental skill levels. The *syllable counting* task required children to count the number of syllables in each word. Because this task assesses sensitivity to syllable structure, it is considered developmentally easier than the other two tasks, both of which assessed sensitivity to phonemic structure. In the *initial consonant matching* task, children heard two words and needed to judge whether they started with the same sound or not. The *final consonant choice* task was considered to be the hardest because it measured the skill expected to be acquired latest. In this task, children heard and repeated a target word and then heard three more words. Children had to select which of those words ended in the same sound as the target. This task was the most difficult because the ability to perform such tasks with final segments develops later than the ability to do them with initial segments (Stanovich, Cunningham, & Cramer, 1984), and because children had to hold four words in a memory buffer. Items for each task are shown in Appendixes A through C.

Specially written software controlled presentation of all stimuli, and the experimenter entered responses directly into the computer. Practice was provided prior to testing on each task. The percentage of correct answers in each task served as the dependent measures. In addition, incorrect answers were recorded and reviewed later for any evidence that children might be implementing a response strategy other than the one required to perform the task correctly. An example of such an alternative response strategy would be if a child responded on the basis of semantic relations between items.

VOT. Children's abilities to label stimuli along a VOT continuum were examined as a way to gauge their general abilities to form phonemic categories when signal structure is available to them. Consonant voicing contrasts are generally more resistant than many other kinds of consonant contrasts to hearing loss (e.g., Boothroyd, 1984), and children with CIs have been found to make these distinctions more readily than they make other consonant contrasts (Giezen, Escudero, & Baker, 2010). The stimuli used here had been used in earlier studies and were found to evoke sharp labeling functions from children, even those with phonological processing problems (Nittrouer, 1999, 2011) or hearing loss (Nittrouer & Burton, 2005). These stimuli are synthetic replicas of *da* and *ta,* consisting of 270-ms vocalic portions. The first formant (F1) has a 40-ms transition at the beginning, going from a starting frequency of 200 Hz to a steady-state frequency of 650 Hz. The second and third formants (F2 and F3) change over the first 70 ms. The second formant starts at 1800 Hz and falls to a steady-state frequency of 1130 Hz. The third formant starts at 3000 Hz and falls to a steady-state frequency of 2500 Hz. The fourth and fifth formants (F4 and F5) are constant throughout the stimuli at 3250 Hz and 3700 Hz, respectively. The fundamental frequency (f0) is 120 Hz for the first 70 ms and then falls linearly through the rest of the stimulus to an ending frequency of 100 Hz. A nine-step continuum was created from these vocalic portions by cutting back the onset of voicing in 5-ms steps from 0 to 40 ms. Before voicing started, no source excited F1, but aspiration noise excited the higher formants. Two 10-ms portions of natural release bursts were used, both from the same male speaker: One came from a production of *da* and one from *ta*. Appending these bursts to the start of the stimuli had the effect of creating a nine-step continuum with effective VOTs of 10 ms to 50 ms.

Practice was provided with natural tokens of *da* and *ta* and then with the endpoint stimuli from the synthetic continuum. These endpoints consisted of the token with the briefest VOT paired with the *da* burst and the token with the longest VOT paired with the *ta* burst. Children needed to respond correctly to nine out of 10 presentations of the endpoints in order to proceed to testing.

Language comprehension. Children's abilities to comprehend spoken language were assessed using the Auditory Comprehension subtest of the Preschool Language Scales, Fourth Edition (Zimmerman, Steiner, & Pond, 2002). This task requires the child to demonstrate an understanding of spoken language by performing specific commands given by the examiner. It is particularly sensitive to children's abilities to comprehend syntactic structures. Standard scores were used as dependent measures.

Vocabulary. Expressive vocabulary was assessed with the Expressive One-Word Picture Vocabulary Test (Brownell, 2000). This task requires the child to provide the words that label a series of pictured items shown one at a time on separate pages. Standard scores were used as dependent measures.

Short-term memory. For the short-term memory task, words were presented over the speaker at a level of 68 dB SPL. Ten lists consisting of the same six words were presented, with the order of words varied across each list randomly by the program. The six words were *ball, coat, dog, ham, pack,* and *rake*. These words and this task have been used with children before (e.g., Nittrouer & Miller, 1999) and are known to be appropriate for children. Words in each list were presented with an onset-to-onset rate of 1 s. After all words were presented, pictures of each item in random order, but not matching that of the audio presentation order, appeared at the top of the computer touch screen. The child's task was to touch each picture in the order heard. As the child touched a picture, it moved down and into place to the right of the picture of the just previously touched word. After all words were touched, the pictures were at the bottom of the screen, in order from left to right according to how the child recalled hearing them. The computer program recorded the child's responses and compared them to the order in which words were actually presented.

Before testing, training was done using the letters *F, H, Q, R, S,* and *Y,* with the same procedures described above for testing. After training on how to do the task with those letter stimuli, the test words were introduced by presenting them over the speaker one at a time and showing them one at a time. All six pictures were then displayed, and the child had to select the picture matching each word presented in isolation to proceed to testing. After testing with the 10 lists, this procedure was repeated. Data were eliminated from the analysis if the child could not respond with perfect accuracy to the six words presented in isolation. The percentage of items out of 60 (10 lists of six words each) for which order was accurately recalled was used as the dependent measure.

Speed of perceptual processing. For a measure of processing speed, we used the Object Naming subtest of the Comprehensive Test of Phonological Processing (Wagner, Torgesen, & Rashotte, 1999). This task consisted of two pages of pictures arranged in four rows of nine pictures each. The child's task was to name the pictures, in order, as quickly as possible. The time required to name all 36 items was obtained from the videotape of the test session, and the sum across the two trials was used as the dependent measure.

Results

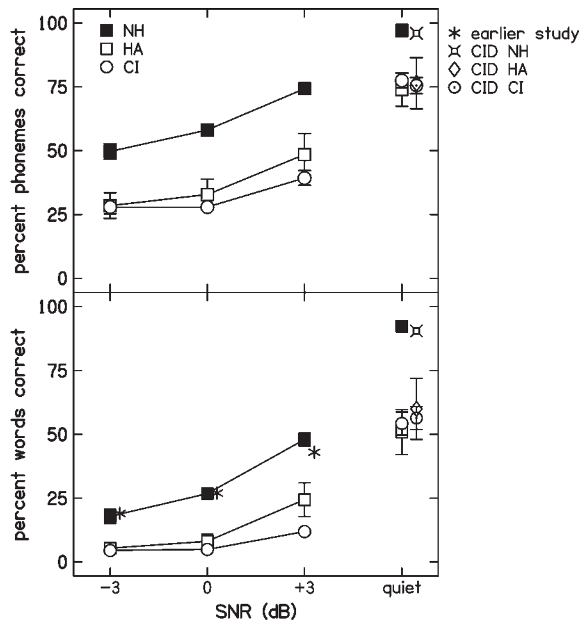
Data were screened to check for homogeneity of variance across groups and for linearity. Arcsine transformations of percentage correct word and phoneme recognition scores were used in inferential tests. An alpha level of .05 was applied, but precise p values are given when $p < .10$. When $p > .10$, results are described as not significant.

Speech Recognition

External and internal validity. Figure 1 shows mean percentage correct recognition for phonemes (top panel) and words (bottom panel) at each SNR and in quiet for each group, with standard errors of the means as error bars. In addition to these group results for the Mackersie et al. (2001) words presented in noise and in quiet, other group results are shown for comparison. For word recognition in noise, results for the same stimuli presented at the same SNRs to 14 typically developing 7-year-olds with NH from Nittrouer et al.'s (2011) study are shown next to mean scores for children with NH from this current study at each SNR. A two-way repeated-measures analysis of variance (ANOVA) on scores for children with NH in the two studies, with study as the between-subjects factor and SNR as the within-subject factor, revealed only a main effect of SNR, $F(2, 62) = 63.26$, $p < .001$. Neither the main effect of study nor the SNR × Study interaction were significant, meaning $p > .10$. This outcome provides reassurance that participants with NH were likely representative of children with NH.

Mean percentage correct recognition scores for phonemes and words presented in quiet are also shown for the CID W-22 word lists. These scores are from the children in the current experiment. For all groups of children, phoneme and word recognition scores appear to be similar

Figure 1. Mean percentage correct recognition for phonemes (top panel) and words (bottom panel), tested at signal-to-noise ratios (SNRs) of −3 dB, 0 dB, and +3 dB, as well as in quiet, for children with normal hearing (NH), hearing aids (HAs), and cochlear implants (CIs). Error bars represent standard errors of the mean (SEM). For mean word recognition (bottom panel), asterisks indicate results from an earlier study (Nittrouer et al., 2011), using the same stimuli for comparison.



for the Mackersie et al. (2001) words and the CID W-22 words, and a series of paired-sample *t* tests carried out on phoneme and word scores for each group separately failed to reveal any statistically significant differences for any group. Thus, the words used in the current test of speech perception in noise are neither harder nor easier than what children are routinely asked to recognize in such tests.

Group differences. We examined potential group differences for recognition of the Mackersie et al. (2001) words in quiet and in noise. For recognition in quiet, we performed one-way ANOVAs with post hoc contrasts. For both phonemes and words, the overall effect of listener group was significant: phonemes, $F(2, 51) = 30.33$, $p < .001$, and words, $F(2, 51) = 31.10$, $p < .001$. In both cases, contrasts between children with NH and children in both of the other two groups were significant ($p < .001$), both with and without Bonferroni adjustments, but contrasts between children with HAs and CIs were not. Thus, one can conclude that children with NH had better recognition in quiet than children with either HAs or CIs. In quiet, recognition was similar for children with hearing loss, regardless of whether they wore HAs and CIs.

The results of two-way repeated-measures ANOVAs performed on scores obtained for words presented in noise are shown in Table 2. Effects of SNR and group membership were both significant, for both phoneme and word scores. The Group × SNR interaction was close to significant for phonemes ($p = .058$) and was significant for words ($p = .028$). For both phoneme and word scores, post hoc contrasts showed differences between children with NH and the two hearing loss groups ($p < .001$), both with and without Bonferroni adjustments. The contrast of scores between children with HAs and CIs was not significant for phoneme scores but was significant for word scores ($p = .040$), without a Bonferroni adjustment; with the adjustment, the contrast was not significant.

Table 2. Results for performance in noise for phoneme and word recognition.

Test	*df*	*F*	*p*	Partial η^2
Phonemes				
SNR	2, 102	42.15	< .001	.453
Group	2, 51	31.21	< .001	.550
SNR × Group	4, 102	2.36	.058	.085
Words				
SNR	2, 102	56.20	< .001	.524
Group	2, 51	60.80	< .001	.705
SNR × Group	4, 102	3.90	.027	.133

Note. SNR = signal-to-noise ratio.

There is one potential flaw in viewing performance strictly according to how well children recognized phonemes and words in noise: Differences were observed across groups for recognition in quiet. Consequently, scores for recognition in noise were influenced not only by the presence of that noise but by general speech recognition abilities as well. To control for the effects of how well children could recognize words in quiet, we computed percentages of phonemes and words recognized correctly in noise for only those phonemes and words recognized correctly in quiet. For phonemes, this approach meant that recognition in noise was matched to recognition in quiet, based on specific words and locations of phonemes within those words. Figure 2 shows group means for these scores, and Table 3 shows results of ANOVAs performed on these conditional phoneme and word scores. Overall, values and patterns of scores in Figure 2 closely resemble those of Figure 1. Statistical outcomes were similar, but the Group × SNR interactions were not significant for these conditional scores. Again, post hoc contrasts showed differences in scores for children with NH and those of the two hearing loss groups ($p < .001$), for both phonemes and words, and again, no significant difference was found between children with HAs and those with CIs for phoneme scores, but for word scores, the contrast was significant ($p = .047$) without a correction for multiple contrasts; with a Bonferroni adjustment, the effect was no longer significant.

Table 3. Results for performance in noise for phoneme and word recognition when that phoneme or word was recognized correctly in quiet.

Test	*df*	*F*	*p*	Partial η^2
Phonemes				
SNR	2, 102	43.54	< .001	.461
Group	2, 51	24.65	< .001	.492
SNR × Group	4, 102	1.73	*ns*	
Words				
SNR	2, 102	56.72	< .001	.527
Group	2, 51	40.37	< .001	.613
SNR × Group	4, 102	1.63	*ns*	

Note. Precise *p* values are shown if they are less than .10; *ns* denotes values greater than .10.

Lexical effects. Mean *j* factors for each group at each SNR are shown in Table 4. These factors were computed only on data from children who had recognition scores between 5% and 95% correct for both phonemes and words. All children with NH had scores within this range at all three SNRs, but some children with hearing loss did not achieve 5% correct recognition at one or more SNRs. These reported *j* factors are similar to those reported by Nittrouer and Boothroyd (1990) for children between 4 and 6 years of age: +3 dB SNR = 2.51 and 0 dB SNR = 2.45. Children were not tested at –3 dB SNR in that study.

Figure 2. Mean percentage correct recognition for only those phonemes (top panel) and words (bottom panel) recognized correctly in quiet, for children with NH, HAs, and CIs. Phonemes and words were tested at SNRs of –3 dB, 0 dB, and +3 dB, as well as in quiet. Error bars represent SEM.

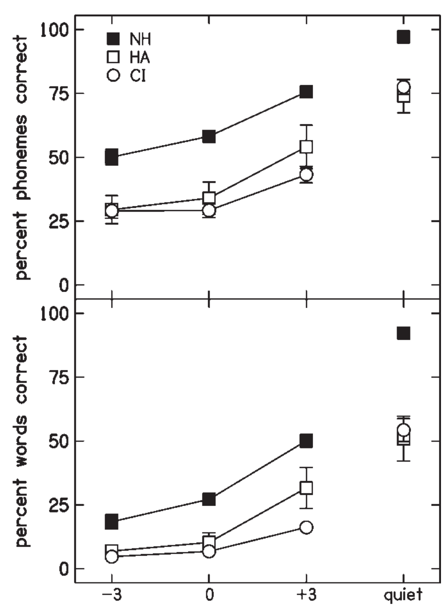


Table 4. Mean *j* factors.

Group	−3 dB SNR			0 dB SNR			+3 dB SNR		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*
Normal hearing	19	2.66	0.35	19	2.49	0.36	19	2.58	0.34
Hearing aids	4	2.56	0.15	5	2.61	0.24	6	2.35	0.25
Cochlear implants	10	2.72	0.45	15	2.50	0.41	20	2.51	0.37

Note. *n* = number of children in each group who had both phoneme and word recognition scores between 5% and 95% so that *j* factors could be computed.

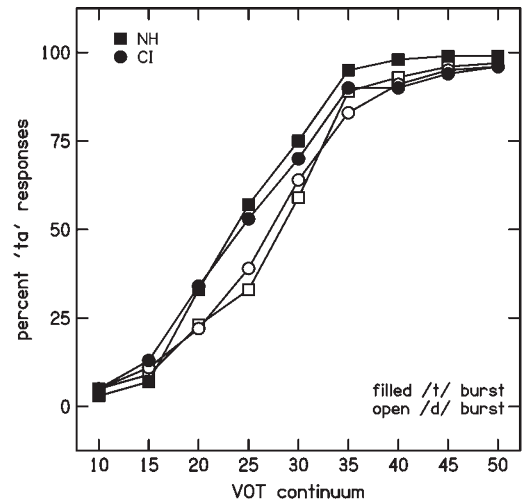
A two-way repeated-measures ANOVA run on the current data showed a significant effect of SNR, $F(2, 58) = 3.15$, $p = .05$, but neither the group effect nor the Group × SNR interaction was significant. Simple effects analyses conducted on these values revealed a significant difference only between *j* factors for –3 dB versus +3 dB SNR ($p = .049$), although the comparison of –3 dB versus 0 dB was close to significant ($p = .058$). The comparison of *j* factors for 0 dB versus +3 dB SNR was not significant; thus, one can conclude that similar lexical effects were found for children in all groups. A slight trend was found toward reduced lexical effects at the poorest SNR.

Summary. Speech recognition diminished as SNR became poorer, and children with hearing loss did not perform as well as children with NH, regardless of whether they wore HAs or CIs. These differences could not be attributed to differences in lexical effects across groups. Significant differences were not observed between children with HAs and CIs for phoneme recognition in noise; however, for word recognition, contrasts reached statistical significance without Bonferroni corrections. Because the group of children with HAs was small, it is possible that a true difference between these groups may exist and was missed in these analyses. A difference in speech perception in noise for children with HAs and CIs could have been predicted because CIs do not preserve all the kinds of signal structure available through HAs. Because of the possibility of a group difference between children with HAs and those with CIs, results from children with HAs were not included in subsequent analyses; instead, analyses focused on potential differences between children with NH and those with CIs.

Other Measures

VOT. Labeling functions for the synthetic VOT stimuli are shown in Figure 3. Functions are very similar in location and slope for children with NH and those with CIs, and *t* tests done on phoneme boundaries and slopes all failed to reveal significant differences between children with NH and those with CIs. Thus, one can conclude that these children with CIs exhibited typical abilities for categorizing speech stimuli that are distinguishable based on the temporal cue of VOT.

Figure 3. Percentage *ta* responses for labeling of synthetic voice-onset time (VOT) stimuli, along a 0- to 50-ms VOT continuum.



General language. The top two rows of Table 5 show mean standard scores, standard deviations, outcomes of *t* tests, and effect sizes in the form of Cohen's *d*s for both auditory comprehension and expressive vocabulary measures. For both measures, children with NH performed better than children with CIs. These differences were significant, with large effect sizes.

Cognitive skills. The middle two rows of Table 5 show mean scores and other statistics for measures of cognitive skills. Children with NH performed better on both tasks than children with CIs. These differences were significant, but with effects of only moderate size.

Phonological awareness. The bottom three rows of Table 5 show mean scores and other statistics for the three measures of phonological awareness. Children with NH performed better than children with CIs on all three, but effects were greater for the two measures of phonemic awareness than the one measure of syllabic awareness. That likely reflects the fact that CIs preserve the amplitude structure affiliated with syllabic structure in the speech signal but are poor at conveying many of the acoustic cues associated with phonemes. In particular, spectral structure arising from the vocal tract filter is lacking.

A review of incorrect responses failed to reveal any evidence that any strategy other than the one designated by the task was used by any child for responding.

Regression Analyses

Separate linear regressions with one predictor variable in each analysis were performed with each of two

Table 5. Means and standard deviations for children with NH and those with CIs, as well as results of *t* tests and Cohen's *d*s for seven measures.

Measure	NH		CIs					
	M	*SD*	*M*	*SD*	*t*	*df*	*p*	Cohen's *d*
General language								
Auditory comprehension	103	10	77	20	5.10	44	< .001	1.64
Expressive vocabulary	110	11	89	18	4.51	44	< .001	1.41
Cognitive skills								
Short-term memory	30	13	23	10	2.06	40	.046	0.60
Rapid serial naming	96	27	128	60	−2.16	42	.037	−0.69
Phonological awareness								
Syllable counting	67	37	46	30	2.06	42	.046	0.62
Initial consonant matching	93	10	64	21	5.40	41	< .001	1.76
Final consonant choice	59	22	14	15	8.21	42	< .001	2.39

Note. Standard scores are shown for both language measures. For short-term memory, the percentages of correct words across the six list positions are shown. For rapid serial naming, time in seconds to name all objects on both pages is shown. For all phonological awareness tasks, the percentages of correct responses are shown. For children with NH, $n = 19$ for all measures, except initial consonant matching, where $n = 18$ because one child became ill during testing. For children with CIs, $n = 27$ for auditory comprehension and expressive vocabulary, $n = 25$ for most other measures because two children (not always the same) were unable to meet criteria for participation, but $n = 23$ for short-term memory because four children were unable to correctly label objects based on auditory-only presentation.

dependent measures of speech recognition in noise. The two measures of speech recognition used in these analyses were the conditional (upon correct recognition in quiet) phoneme and word recognition scores. These scores were selected for use on the basis of principled grounds (they seem most valid), but in fact outcomes were no different when the nonconditional percentage-correct scores were used. The mean across the three SNRs for each measure (phoneme and word recognition) was computed for each child. The predictor variables used were each of the seven measures shown in Table 5. These linear regressions were performed with scores for children with NH and those with CIs combined, as well as for each group separately. Standardized beta coefficients are shown in Table 6.

When we looked at scores for all children, we found that the predictor variables that explained significant proportions of variance in phoneme and word recognition were the language measures and the measures of phonemic awareness. The highest standardized beta coefficient obtained was for word recognition and the final consonant choice task ($\beta = .782$). Short-term memory also explained a significant proportion of variance for word recognition ($\beta = .405$), but the magnitude of that effect was smaller than for any general language or phonemic awareness measure. When we examined beta coefficients for each group separately, we noted that the only one that explained a significant proportion of variance for either group was the one for the initial consonant matching task and word recognition for children with NH ($\beta = .594$).

Table 6. Standardized beta coefficients for each predictor variable and the dependent variables of phoneme recognition and word recognition.

Measure	General language		Cognitive skills		Phonological awareness		
	Auditory comp.	Expressive vocabulary	Short-term memory	Rapid naming	Syllable counting	Initial consonant	Final consonant
Phoneme recognition							
All children	.562**	.533**	.390	−.209	.286	.546**	.715**
NH	.391	.239	.332	.085	−.016	.341	.355
CI	.114	.146	.306	.097	.106	.018	.223
Word recognition							
All children	.612**	.577**	.405*	−.274	.290	.625**	.782**
NH	.408	.252	.294	.074	.053	.594*	.386
CI	.130	.182	.336	.006	.047	.043	.254

Note. comp. = comprehension.

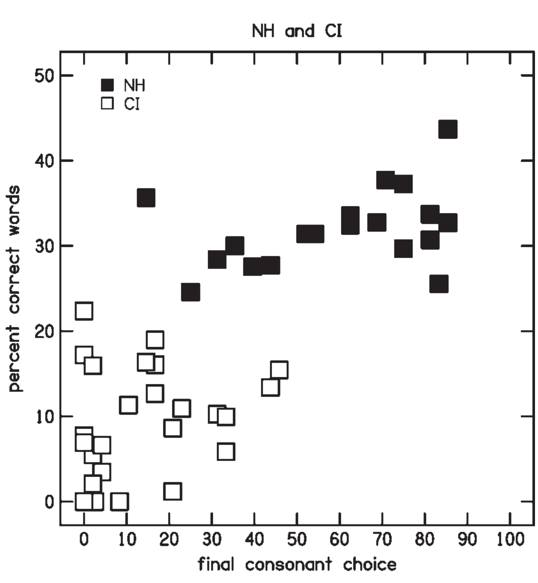
*$p < .01$. **$p < .001$.

The pattern of finding significant beta coefficients when children in both groups are considered together, but finding no significant coefficients when they are considered separately, means that the two groups differ in performance for both variables (predictor and dependent), but there was not a continuous effect within the groups. Figure 4 illustrates this pattern for word recognition and final consonant choice, the analysis that resulted in the largest beta coefficient when all children were included. The pattern found here is one of two distinct groups defined by scores on each of the measures but no indication of a relationship between scores for either group. In fact, several children with NH performed similarly to children with CIs on the final consonant choice task but were more accurate at recognizing words in noise.

Factors Affecting Performance of Children With CIs

We again used the conditional phoneme and word recognition scores as dependent variables to examine potential effects of factors related to hearing loss and cochlear implantation in the children with CIs. First, we obtained standardized beta coefficients using age of identification, age at first implantation, and age at second implantation (for the 18 children who had two implants) as predictor variables. Of these, only age at first implantation revealed significant effects for phoneme recognition (β = .424, t = 2.341, p = .027), but not for word recognition, (β = .362, t = 1.945, p = .063). Thus, being younger at the time of the first implant was associated with better outcomes for speech recognition in noise: Roughly 18% of the variance in phoneme recognition scores was explained by age at first implant.

Figure 4. Scatter plot showing means across three SNRs as a function of percentage words correct on the final consonant choice task for children with NH and children with CIs.



Next, we examined the effects of having one or two implants and of having had or not had some period of bimodal experience. Only one child was still wearing an HA and a CI at the time of testing. That child was not included in the analyses of one or two implants because she did not fit neatly into either group. Mean conditional phoneme and word recognition scores are shown in Table 7, and t tests were performed on these scores. The effect of having one or two implants was not significant (p > .10). The effect of having had some period of bimodal experience was not significant either; however, a trend toward better word recognition (t = 2.038, p = .052) was found for the group with some period of bimodal experience.

Finally, we examined scores for phoneme and word recognition in quiet for children with CIs. Because their performance was not close to 100% accurate in quiet, as was the case for children with NH, the abilities of children with CIs to recognize speech in noise were highly constrained by their abilities to recognize speech in quiet. Thus, it seemed important to determine what explained those quiet recognition abilities. Stepwise linear regressions were performed, with the percentages of phonemes and words recognized correctly in quiet serving as dependent variables. Predictor variables included the seven measures that have been considered thus far (shown in Table 5), as well as age of first implantation, because that was found to explain a significant proportion of variance for phoneme recognition in noise.

The results of these analyses are shown in Table 8. For both phoneme and word recognition in quiet, expressive vocabulary scores explained a large proportion of variance, such that the better a child's vocabulary was, the better the child's recognition in quiet was. Scores on the initial consonant matching task were also found to

Table 7. Means and standard deviations of conditional word and phoneme recognition scores for children with CIs, as a function of number of implants and whether they had some bimodal experience.

	Number of implants			
	One (n = 8)		Two (n = 18)	
Measure	*M*	*SD*	*M*	*SD*
Phoneme recognition	33.5	9.8	32.8	12.7
Word recognition	9.7	6.4	8.3	5.8
	Bimodal experience			
	Yes (n = 12)		No (n = 14)	
Phoneme recognition	35.0	10.0	31.3	13.1
Word recognition	10.8	5.8	7.0	5.5

Table 8. Standardized beta coefficients of predictor variables explaining significant portions of variance for phoneme and word recognition in quiet for children with CIs, derived from stepwise linear regression with F to enter = .05.

Measure	Standardized coefficient	t	p	R^2 for model
Phoneme recognition				
Expressive vocabulary	.705	3.57	.002	.407
Initial consonant matching	−.414	−2.10	.050	
Word recognition				
Expressive vocabulary	.709	3.62	.002	.418
Initial consonant matching	−.445	−2.27	.035	

explain significant proportions of variance in phoneme and word recognition scores. However, the interesting aspect of this result is that the direction of relationship was opposite to what would be expected: Higher scores on the initial consonant task were associated with lower scores on phoneme and word recognition in quiet.

Discussion

We undertook this study to examine factors that were considered likely candidates to help explain speech recognition in noise by children who use CIs. We fully anticipated before the study was conducted that children with CIs would perform more poorly than children with NH on measures of speech recognition in noise, and that was indeed found to be the case. However, only small Group × SNR interactions were obtained, and only when scores not contingent on performance in quiet were used in statistical analyses. Those patterns of result suggest that noise effects on recognition were either similar or only slightly greater for children with CIs than for children with NH. This finding could be interpreted as indicating that—at least for children—the limitations on the kind of signal delivered by CIs take the greatest toll on speech recognition in quiet. Adding noise to the signal appears to have comparable effects for all children, regardless of hearing status.

That situation differs from the one presumed to exist for adults. It is commonly believed that noise has more deleterious consequences for implant users than for listeners with NH (e.g., Carroll, Tiaden, & Zeng, 2011), but perusal of the literature fails to provide strong evidence of this position. Few studies have quantified phoneme or word recognition scores in various levels of noise for adults with NH and those with CIs (cf. Zeng & Galvin, 1999). One report that gave scores derived from real words showed close to a 40-percentage-point drop in phoneme recognition scores between 25 dB SNR (which is effectively listening in quiet) and 0 dB SNR for implant users and listeners with NH alike (Hochberg, Boothroyd, Weiss, & Hellman, 1992). Those recognition scores match what is typically found for listeners with NH (e.g., Boothroyd & Nittrouer, 1988) and for listeners with CIs (e.g., Friesen, Shannon, Baskent, & Wang, 2001) at those noise levels. Consequently, it could be that the processing limitations imposed by CIs have their primary effects on recognition in quiet even for adults, and noise effects are consistent in magnitude for listeners with NH and those who use CIs. Be that as it may for adults, the current study suggests that it is likely the case for children.

Of course, these results spark the question of why children with CIs are so much poorer at recognizing speech than children with NH, even in quiet. Some of the explanation must surely be that implants provide only a sparse signal representation, lacking many of the acoustic properties, especially spectral ones, traditionally thought to underlie phonetic perception. At the same time, children with CIs demonstrate a variety of language deficits, including those of vocabulary, phonological awareness, and comprehension of language structures. Although the source of those difficulties may very well be traced to the impoverished signals provided through implants, the language and phonological deficits would nonetheless be expected to feed back and have negative effects of their own on speech recognition. As Ahissar (2007) suggested, highly refined language knowledge allows listeners to make predictions about the structure that is likely to be in the signal; conversely, the absence of such knowledge inhibits recognition. This phenomenon presumably would be just as applicable to perception in quiet as in noise, especially if the signal is degraded due to processing limitations.

Nonetheless, children with CIs performed significantly more poorly than children with NH when it came to recognizing speech in noise. This difference between groups was strongly related to group differences in language and phonological awareness measures, but no evidence was found to suggest that recognition scores in noise were explained by the language and/or phonological awareness measures for children with CIs: No within-groups effects were found. Only one measure (initial consonant matching) was found to explain a significant proportion of variance for recognition in noise, but only for the children with NH.

Given the trends observed in this study, one can conclude that the difficulties in recognizing speech in noise experienced by children with CIs differ from those of children with dyslexia. Children in the latter group recognized speech perfectly well in quiet and exhibited problems only for recognition in noise. Thus, there is a clear and disproportionately large effect of noise on recognition abilities for children with dyslexia, compared to that experienced by children with NH and without dyslexia. Children with CIs showed deficits for speech recognition in quiet and an effect of noise similar in magnitude to what was observed for children with NH. The implication of these findings is that improved processing strategies for CIs, if and when they are developed, should be expected to have positive effects on recognition in both quiet and noise for children with CIs. In fact, if processing strategies that preserve phonemically relevant spectral structure in the signal and/or temporal fine structure were to be developed, children with CIs could reasonably be expected to acquire better language and phonological skills as well. That prediction follows from the fact that results from the VOT continua revealed that these children are perfectly capable of forming well-defined phonological categories, as long as they have access to the signal properties on which those categories are based. Until such time as improved processing strategies can be implemented in CIs, teachers and clinicians will need to continue providing extra support to these children to help them acquire the language and phonological skills they need to succeed in academic settings. Care should be given to maximize signal-to-noise levels for these children, as well.

Acknowledgments

This work was supported by Grant R01 DC006237 from the National Institute on Deafness and Other Communication Disorders to Susan Nittrouer. We thank Daniel Burry, Joanna Lowenstein, Eric Tarr, and Puisan Wong for help with stimulus development and data collection, and Caitlin Rice for help with manuscript preparation.

References

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences, 11,* 458–465.

Ahissar, M., Lubin, Y., Putter-Katz, H., & Banai, K. (2006). Dyslexia and the failure to form a perceptual anchor. *Nature Neuroscience, 9,* 1558–1564.

Baer, T., Moore, B. C., & Gatehouse, S. (1993). Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: Effects on intelligibility, quality, and response times. *Journal of Rehabilitation Research & Development, 30,* 49–72.

Bernstein, J. G., & Brungart, D. S. (2011). Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio. *The Journal of the Acoustical Society of America, 130,* 473–488.

Boothroyd, A. (1968). Statistical theory of the speech discrimination score. *The Journal of the Acoustical Society of America, 43,* 362–367.

Boothroyd, A. (1984). Auditory perception of speech contrasts by subjects with sensorineural hearing loss. *Journal of Speech and Hearing Research, 27,* 134–144.

Boothroyd, A., Mulhearn, B., Gong, J., & Ostroff, J. (1996). Effects of spectral smearing on phoneme and word recognition. *The Journal of the Acoustical Society of America, 100,* 1807–1818.

Boothroyd, A., & Nittrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *The Journal of the Acoustical Society of America, 84,* 101–114.

Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *The Journal of the Acoustical Society of America, 106,* 2074–2085.

Brady, S., Shankweiler, D., & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35,* 345–367.

Breier, J. I., Gray, L., Fletcher, J. M., Diehl, R. L., Klaas, P., Foorman, B. R., & Molis, M. R. (2001). Perception of voice and tone onset time continua in children with dyslexia with and without attention deficit/hyperactivity disorder. *Journal of Experimental Child Psychology, 80,* 245–270.

Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test* (3rd ed.). Novato, CA: Academic Therapy.

Burkholder, R. A., & Pisoni, D. B. (2003). Speech timing and working memory in profoundly deaf children after cochlear implantation. *Journal of Experimental Child Psychology, 85,* 63–88.

Carroll, J., Tiaden, S., & Zeng, F. G. (2011). Fundamental frequency is critical to speech perception in noise in combined acoustic and electric hearing. *The Journal of the Acoustical Society of America, 130,* 2054–2062.

Cleary, M., Pisoni, D. B., & Geers, A. E. (2001). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear and Hearing, 22,* 395–411.

Cutler, A., Garcia Lecumberri, M. L., & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *The Journal of the Acoustical Society of America, 124,* 1264–1268.

Davidson, L. S., Geers, A. E., Blamey, P. J., Tobey, E. A., & Brenner, C. A. (2011). Factors contributing to speech perception scores in long-term pediatric cochlear implant users. *Ear and Hearing, 32,* 19S–26S.

Firszt, J. B., Holden, L. K., Skinner, M. W., Tobey, E. A., Peterson, A., Gaggl, W., . . . Wackym, P. A. (2004). Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems. *Ear and Hearing, 25,* 375–387.

Flege, J. E., MacKay, I. R., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America, 106,* 2973–2987.

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America, 110,* 1150–1163.

Fu, Q. J., Shannon, R. V., & Wang, X. (1998). Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing. *The Journal of the Acoustical Society of America, 104,* 3586–3596.

Gaab, N., Gabrieli, J. D., Deutsch, G. K., Tallal, P., & Temple, E. (2007). Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: An fMRI study. *Restorative Neurology and Neuroscience, 25,* 295–310.

Geers, A. E., & Hayes, H. (2011). Reading, writing, and phonological processing skills of adolescents with 10 or more years of cochlear implant experience. *Ear and Hearing, 32,* 49S–59S.

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

Glasberg, B. R., & Moore, B. C. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America, 79,* 1020–1033.

Godfrey, J. J., Syrdal-Lasky, A. K., Millay, K. K., & Knox, C. M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology, 32,* 401–424.

Hayes, H., Geers, A. E., Treiman, R., & Moog, J. S. (2009). Receptive vocabulary development in deaf children with cochlear implants: Achievement in an intensive auditory-oral educational setting. *Ear and Hearing, 30,* 128–135.

Hirsh, I. J., Reynolds, E. G., & Joseph, M. (1954). Intelligibility of different speech materials. *The Journal of the Acoustical Society of America, 26,* 530–538.

Hochberg, I., Boothroyd, A., Weiss, M., & Hellman, S. (1992). Effects of noise and noise suppression on speech perception by cochlear implant users. *Ear and Hearing, 13,* 263–271.

James, D., Rajput, K., Brinton, J., & Goswami, U. (2009). Orthographic influences, vocabulary development, and phonological awareness in deaf children who use cochlear implants. *Applied Psycholinguistics, 30,* 659–684.

Johnson, C., & Goswami, U. (2010). Phonological awareness, vocabulary, and reading in deaf children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 53,* 237–261.

Leek, M. R., Dorman, M. F., & Summerfield, Q. (1987). Minimum spectral contrast for vowel identification by normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 81,* 148–154.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 18866–18869.

Mackersie, C. L., Boothroyd, A., & Minniear, D. (2001). Evaluation of the Computer-Assisted Speech Perception Assessment Test (CASPA). *Journal of the American Academy of Audiology, 12,* 390–396.

Manis, F. R., McBride-Chang, C., Seidenberg, M. S., Keating, P., Doi, L. M., Munson, B., & Petersen, A. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Child Psychology, 66,* 211–235.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996, January 5). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271,* 77–81.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology, 41,* 329–335.

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N. Y., Quittner, A. L., . . . CDaCI Investigative Team. (2010). Spoken language development in children following cochlear implantation. *Journal of the American Medical Association, 303,* 1498–1506.

Nittrouer, S. (1999). Do temporal processing deficits cause phonological processing problems? *Journal of Speech, Language, and Hearing Research, 42,* 925–942.

Nittrouer, S. (2010). *Early development of children with hearing loss*. San Diego, CA: Plural Publishing.

Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America, 87,* 2705–2715.

Nittrouer, S., & Burton, L. T. (2005). The role of early language experience in the development of speech perception and phonological processing abilities: Evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. *Journal of Communication Disorders, 38,* 29–63.

Nittrouer, S., & Miller, M. E. (1999). The development of phonemic coding strategies for serial recall. *Applied Psycholinguistics, 20,* 563–588.

Nittrouer, S., Shune, S., & Lowenstein, J. H. (2011). What is the deficit in phonological processing deficits: Auditory sensitivity, masking, or category formation? *Journal of Experimental Child Psychology, 108,* 762–785.

Pinet, M., & Iverson, P. (2010). Talker–listener accent interactions in speech-in-noise recognition: Effects of prosodic manipulation as a function of language experience. *The Journal of the Acoustical Society of America, 128,* 1357–1365.

Pisoni, D. B., Kronenberger, W. G., Roman, A. S., & Geers, A. E. (2011). Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear and Hearing, 32,* 60S–74S.

Pollack, I., Rubenstein, H., & Decker, L. (1959). Intelligibility of known and unknown message sets. *The Journal of the Acoustical Society of America, 31,* 273–279.

Revoile, S. G., Pickett, J. M., & Kozma-Spyteck, L. (1991). Spectral cues to perception of /d, n, l/ by normal- and impaired-hearing listeners. *The Journal of the Acoustical Society of America, 90,* 787–798.

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics, 27,* 465–485.

Rubinstein, J. T. (2004). How cochlear implants encode speech. *Current Opinion in Otolaryngology & Head and Neck Surgery, 12,* 444–448.

Serniclaes, W., Ventura, P., Morais, J., & Kolinsky, R. (2005). Categorical perception of speech sounds in illiterate adults. *Cognition, 98,* B35–B44.

Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 38,* 175–190.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language, 9,* 182–198.

Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. In P. Tallal (Ed.), *Annals of the New York Academy of Sciences: Vol. 682. Temporal information processing in the* *central nervous system: Special reference to dyslexia and dysphasia* (pp. 27–47). New York, NY: New York Academy of Sciences.

Tallal, P., & Piercy, M. (1973, February 16). Defects of nonverbal auditory perception in children with developmental aphasia. *Nature, 241,* 468–469.

Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia, 12,* 83–93.

Tomblin, J. B., Spencer, L., Flock, S., Tyler, R., & Gantz, B. (1999). A comparison of language achievement in children with cochlear implants and children using hearing aids. *Journal of Speech, Language, and Hearing Research, 42,* 497–511.

Vance, M., & Martindale, N. (2011). Assessing speech perception in children with language difficulties: Effects of background noise and phonetic contrast. *International Journal of Speech-Language Pathology, 14,* 48–58.

von Hapsburg, D., Champlin, C. A., & Shetty, S. R. (2004). Reception thresholds for sentences in bilingual (Spanish/English) and monolingual (English) listeners. *Journal of the American Academy of Audiology, 15,* 88–98.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *The Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.

Zeng, F. G., & Galvin, J. (1999). Amplitude mapping and phoneme recognition in cochlear implant listeners. *Ear and Hearing, 20,* 60–74.

Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F. X., & Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 14110–14115.

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2011). Noise on, voicing off: Speech perception deficits in children with specific language impairment. *Journal of Experimental Child Psychology, 110,* 362–372.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool Language Scale, Fourth Edition*. San Antonio, TX: The Psychological Corporation.

Appendix A. Syllable counting.

Pre-Practice Examples With Natural Voice

A. The subject's name
B. Either a sibling's or a pet's name
C. The words *cat* (1), *catnap* (2), and *catnapping* (3)

Practice

A. but	D. tell	G. doll	J. top
B. butter	E. telling	H. dolly	K. water
C. butterfly	F. telephone	I. lollipop	L. elephant

****Discontinue after 6 consecutive errors.**

Test Trials

1. popsicle (3)	25. grab (1)
2. dinner (2)	26. rectangle (3)
3. penny (2)	27. bird (1)
4. house (1)	28. location (3)
5. valentine (3)	29. arrive (2)
6. open (2)	30. cap (1)
7. box (1)	31. cowboy (2)
8. cook (1)	32. beach (1)
9. birthday (2)	33. forecast (2)
10. president (3)	34. daylight (2)
11. bicycle (3)	35. calendar (3)
12. typewriter (3)	36. leaf (1)
13. green (1)	37. telescope (3)
14. gasoline (3)	38. decorate (3)
15. chicken (2)	39. crocodile (3)
16. letter (2)	40. grandson (2)
17. jump (1)	41. vine (1)
18. morning (2)	42. cold (1)
19. dog (1)	43. teacher (2)
20. monkey (2)	44. memorize (3)
21. anything (3)	45. drive (1)
22. wind (1)	46. useful (2)
23. nobody (3)	47. weekend (2)
24. boat (1)	48. factory (3)

Appendix B. Initial consonant same/different.

Practice Examples

1. bark	barn	4. pet	pack
2. jump	shirt	5. blue	bag
3. mat	cap	6. star	clown

****Discontinue after 6 consecutive incorrect same items.**

Test Trials

1. leap	lip*	25. peel	pat*
2. key	kite*	26. tile	mask
3. crumb	drip	27. note	wheel
4. date	bag	28. meat	lace
5. gate	gum*	29. soap	salt*
6. sky	sleep*	30. day	box
7. grape	glue*	31. wash	vine
8. king	dime	32. zip	zoo*
9. dark	pet	33. stick	slide*
10. toes	tip*	34. plum	price*
11. class	swing	35. win	well*
12. web	man	36. pear	pen*
13. tree	star	37. soup	light
14. milk	moon*	38. frog	brush
15. pin	boat	39. fist	sap
16. claw	crib*	40. met	map*
17. lock	pail	41. house	heel*
18. bit	girl	42. leg	lock*
19. foot	pan	43. prize	stair
20. drum	flag	44. rain	kid
21. bone	bud*	45. sled	stick*
22. fun	fan*	46. sun	bin
23. rug	rag*	47. jeep	jug*
24. can	pit	48. duck	door*

***Same**

Appendix C. Final consonant choice.

Practice Examples

1. rib	<u>mob</u> phone heat	4. lamp	<u>tip</u> rock juice
2. stove	<u>cave</u> hose stamp	5. fist	<u>hat</u> knob stem
3. hoof	<u>tough</u> shed cop	6. head	<u>rod</u> hem fork

****Discontinue after 6 consecutive errors.**

Test Trials

1. truck	wave trust <u>bike</u>	25. desk	<u>lock</u> path tube
2. duck	bath song <u>rake</u>	26. home	<u>drum</u> prince mouth
3. mud	mug <u>crowd</u> dot	27. leaf	<u>roof</u> suit leak
4. sand	sash <u>kid</u> flute	28. thumb	tub jug <u>cream</u>
5. flag	cook <u>rug</u> step	29. barn	tag night <u>pin</u>
6. car	foot can <u>stair</u>	30. doll	<u>wheel</u> pig beef
7. comb	cob <u>room</u> drip	31. train	grade <u>van</u> cape
8. boat	<u>skate</u> frog bone	32. bear	<u>shore</u> clown rat
9. house	<u>kiss</u> mall dream	33. pan	<u>skin</u> grass beach
10. cup	<u>lip</u> trash plate	34. hand	hail run <u>lid</u>
11. meat	sock <u>date</u> camp	35. pole	<u>mail</u> poke land
12. worm	price <u>team</u> soup	36. ball	<u>pool</u> clip steak
13. hook	<u>neck</u> mop weed	37. park	bed <u>lake</u> crown
14. rain	<u>yawn</u> thief sled	38. gum	shoe <u>lamb</u> gust
15. horse	lunch bag <u>Ice</u>	39. vest	<u>cat</u> star mess
16. chair	slide <u>deer</u> chain	40. cough	log dough <u>knife</u>
17. kite	mouse <u>bat</u> grape	41. wrist	<u>throat</u> risk store
18. crib	<u>job</u> hair wish	42. bug	bus <u>leg</u> rope
19. fish	shop gym <u>brush</u>	43. door	dorm <u>pear</u> food
20. hill	moon hip <u>bowl</u>	44. nose	<u>maze</u> goose zoo
21. hive	light hike <u>glove</u>	45. nail	<u>bill</u> voice chef
22. milk	tail <u>block</u> mitt	46. dress	<u>rice</u> tape noise
23. ant	<u>gate</u> school fan	47. box	book <u>face</u> mask
24. dime	note cube <u>broom</u>	48. spoon	cheese <u>fin</u> back