

Speech perception of sine-wave signals by children with cochlear implants

Susan Nittrouer,^{a)} Jamie Kuess, and Joanna H. Lowenstein

Department of Otolaryngology, The Ohio State University, 915 Olentangy River Road, Suite 4000,
Columbus, Ohio 43212, USA

(Received 14 October 2014; revised 30 March 2015; accepted 8 April 2015)

Children need to discover linguistically meaningful structures in the acoustic speech signal. Being attentive to recurring, time-varying formant patterns helps in that process. However, that kind of acoustic structure may not be available to children with cochlear implants (CIs), thus hindering development. The major goal of this study was to examine whether children with CIs are as sensitive to time-varying formant structure as children with normal hearing (NH) by asking them to recognize sine-wave speech. The same materials were presented as speech in noise, as well, to evaluate whether any group differences might simply reflect general perceptual deficits on the part of children with CIs. Vocabulary knowledge, phonemic awareness, and “top-down” language effects were all also assessed. Finally, treatment factors were examined as possible predictors of outcomes. Results showed that children with CIs were as accurate as children with NH at recognizing sine-wave speech, but poorer at recognizing speech in noise. Phonemic awareness was related to that recognition. Top-down effects were similar across groups. Having had a period of bimodal stimulation near the time of receiving a first CI facilitated these effects. Results suggest that children with CIs have access to the important time-varying structure of vocal-tract formants.

© 2015 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4919316>]

[MSS]

Pages: 2811–2822

I. INTRODUCTION

By now, it is well accepted that young children acquire knowledge about linguistic structure and function from “the outside in.” That is, the earliest unit of linguistic organization for the child appears to be the word, or indivisible phrase (e.g., *all gone*). Children somehow discover these lexical units in the speech they hear, and those speech signals come to them as largely unparsed utterances. Where language is concerned, a child’s first developmental hurdle involves finding word-sized units within these unanalyzed signals; the second is finding phonemic structure within those words. The question may legitimately be asked of how children ever manage to accomplish these seemingly formidable tasks.

One suggestion that has been offered in answer to that question stems from the idea that children attend to the “spectral skeletons” of the speech signal (Nittrouer *et al.*, 2009). This structure is defined as the rather slowly shifting spectral patterns associated with continuously changing vocal-tract configurations; it primarily consists of time-varying patterns of change in the first two or three formants. The specific proposal is that in listening to speech, young children gradually learn to recognize recurring patterns of formant change in the ongoing speech stream (e.g., Nittrouer, 2006). These recurring spectral patterns can be perceptually isolated from the ongoing signal to form rudimentary lexical representations. As children acquire increasing numbers of these lexical representations, they start to turn their attention to the acoustic details associated with

those time-varying patterns of formant change; they begin to attend to short, spectrally discrete sections of the signal that have traditionally been termed “acoustic cues.” It is attention to that level of acoustic detail that provokes the acquisition of phonological (especially phonemic) sensitivity, which in turn leads to the refinement of structure within the child’s lexicon (e.g., Beckman and Edwards, 2000; Ferguson and Farwell, 1975). Gradually towards the end of the first decade of life, the lexicon appears to become *re-organized* with the broad spectral patterns described above being replaced with phonemic structure (e.g., Storkel, 2002; Walley *et al.*, 2003).

Early research on human speech perception focused almost entirely on acoustic cues. The major experimental paradigm used in that early work included sets of stimuli that were identical in all aspects of acoustic structure, except one. The setting of that one piece of structure—or cue—would be manipulated along an acoustic continuum, going from a value appropriate for one phoneme to a value appropriate for another phoneme that formed a minimal pair with the first. The collective purpose of this line of investigation was to compile inventories of cues that define all the phonemic categories within a language, mostly to meet the related goals of developing synthetic speech synthesis and automatic speech recognition (Liberman, 1996).

Of course, that line of investigation was based on the notion that speech perception proceeds by listeners harvesting acoustic cues from the signal, and using them to recover the strings of phonemic segments comprising that signal. In the early days, that view of speech perception was rarely questioned; after all, it fit with the impressions of investigators conducting the work, who were all highly literate adults. Accordingly, when it was observed that infants could

^{a)}Electronic mail: nittrouer.1@osu.edu

discriminate syllables differing by phonemes that formed minimal contrasts, the conclusion was readily reached that infants must recognize phonemic units when they listen to speech, just as it was presumed adults do. Humans must be born with a specialized “phonetic module,” it was reasoned (e.g., Werker, 1991).

Eventually, however, scientists began to uncover findings that would challenge the prevailing perspective of speech perception (i.e., one largely involving listeners harvesting acoustic cues and using them to recover strings of phonemes). Where adults were concerned, a major challenge arose when it was shown that they were able to recognize speech fairly well, even when acoustic cues were mostly eliminated. First, there were experiments with sine-wave replicas of speech, the method to be used in the current experiment. In this signal processing technique, the time-varying frequencies of the first two or three formants are extracted and replaced with sine waves; other components of the speech signal are largely absent. Remez *et al.* (1981) showed that adults were able to repeat sentences presented in this highly degraded form. Although earlier studies had demonstrated that listeners can recognize distorted speech signals, as in the seminal work of Licklider and Pollock (1948) with infinitely peak-clipped signals, this was the first demonstration that listeners could do so when acoustic cues were deliberately eliminated. A few years later, Shannon *et al.* (1995) reported a similar finding when acoustic cues were eliminated by processing speech so that only amplitude structure in a few spectral channels was preserved. These demonstrations that acoustic cues were not essential to speech perception raised questions regarding what the critical elements actually are in this perception.

At the same time that those investigations were being reported, developmental psycholinguists were finding that children do not seem to be as aware of word-internal phonemic structure as would be predicted by the idea of an innate phonetic module. For example, in 1982, Treiman and Breaux demonstrated that whereas adults judge similarity between syllables based on the number of shared phonemes, 4-yr-olds base judgments on overall syllable shape, a quality that is more properly seen as related to acoustic structure. Following from that, Walley *et al.* (1986) examined how many phonemes need to be shared between nonsense disyllables in order for them to be judged as similar by children in kindergarten or second grade. Their results showed that the second-grade children were able to recognize as similar disyllables with one shared phoneme, but the kindergarten children were quite poor at doing so; three shared phonemes were required before they were as accurate at judging similarity between pairs of disyllables as the second-grade children. These sorts of results conflicted with reports coming from scientists studying discrimination of syllables by infants, who claimed that infants have phonemic representations (e.g., Jusczyk, 1995; Kuhl, 1987). According to the studies with school-age children (Treiman and Breaux, 1982; Walley *et al.*, 1986), sensitivity to phonemic structure is continuing to emerge through the second half of the first decade of life. Studies by others supported

that position by showing that children continue to judge and categorize linguistically meaningful signals based on global structure through early childhood (e.g., Charles-Luce and Luce, 1990). Although not well specified in those reports, the slowly changing spectral patterns of formants certainly seem to match those authors’ description of global structure.

Empirical evidence of the importance of this global spectral structure comes from studies comparing sentence recognition by adults and children of sine-wave and noise-vocoded signals. For example, Nittrouer and Lowenstein (2010) asked adults and children of three ages (7-, 5-, and 3-years-old) to repeat five-word sentences from the Hearing in Noise Test, or HINT (Nilsson *et al.*, 1994), presented as either sine-wave or noise-vocoded stimuli. Although both of these forms of signal processing degrade the spectral representation of the signal overall, and eliminate most of the kinds of acoustic structure typically termed cues, the kinds of signals that result from each processing strategy differ. In particular, sine-wave signals are especially good at preserving time-varying patterns of formant change; noise-vocoded signals are very poor at preserving that kind of signal structure. This difference can be seen in Fig. 1, where a sine-wave version (middle panel) and a 4-channel noise-vocoded version (bottom panel) of the same sentence are shown. If the hypothesis is accurate that young children attend particularly strongly to the time-varying patterns of formant change, it could be predicted they would perform disproportionately better with sine-wave than with noise-vocoded sentences. Although it can be difficult to compare results across these processing conditions, two kinds of evidence from the Nittrouer and Lowenstein study supported the hypothesis. First, children performed more similarly to adults for the sine-wave than for the noise-vocoded signals: for example, 3-year-olds scored 63.3 percentage points worse than adults with these noise-vocoded signals, but only 23.0 percentage points worse with the sine-wave stimuli. Furthermore, the difference in performance between conditions was much greater for children than for adults. Again using 3-year-olds for comparison, it was found that adults showed an 18.8 percentage point difference between scores for the sine-wave and noise-vocoded stimuli, whereas 3-year-olds showed a 59.2 percentage point difference. Thus, children showed a disproportionately greater benefit than adults from having time-varying formant structure available to them.

Given the evident value of this kind of structure for language learning, concern can be raised about whether or not it is available to children with hearing loss, especially those who receive cochlear implants (CIs). To a first approximation, the signal processing of CIs can be viewed as implementing the same techniques as methods used in noise vocoding: the spectrum of speech is divided into a few spectral channels, amplitude envelopes from each of those channels are recovered, and then those envelopes are presented to listeners with no other spectral detail. Beyond that, the quality of the signal must be presumed to be more degraded for children with CIs, compared to children with normal hearing (NH) listening to noise-vocoded speech. Signals from separate channels spread rather great distances along the basilar membrane with

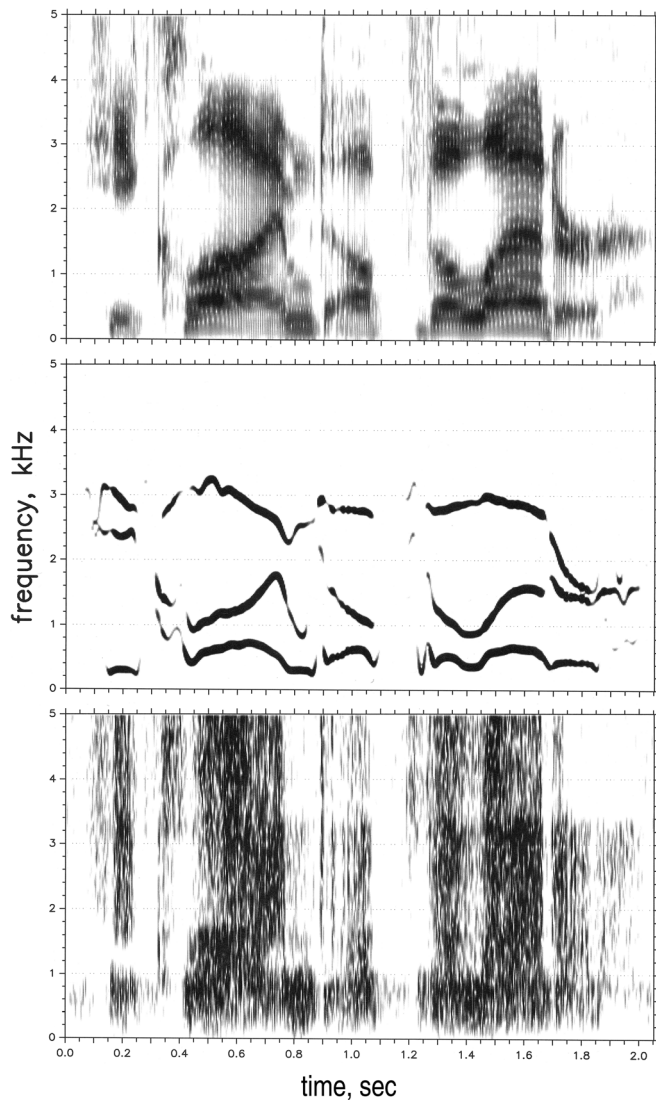


FIG. 1. Spectrograms of the sentence “He climbed up the ladder” from Nittrouer and Lowenstein (2010). The top panel shows the natural production, the middle panel shows a sine-wave version, and the bottom panel shows a 4-channel noise-vocoded version.

CIs stimulation, and regions of neuronal loss can mean that sections of the signal are not transmitted up the auditory system. Due to these constraints, children with CIs would be predicted to have diminished access to time-varying formant structure, compared to children with NH, and the magnitude of that deficit might not be well predicted by results of children with NH listening to noise-vocoded speech.

Another reason to suspect that children with CIs may not be able to recognize time-varying formant structure as well as children with NH is that they have disproportionately more difficulty with one other kind of degraded signal, which is speech in noise. Children with CIs have been found to be poorer at recognizing speech in noise than children with NH (e.g., Caldwell and Nittrouer, 2013; Nittrouer *et al.*, 2013; Smiljanic and Sladen, 2013). Consequently, it was considered possible before this study was conducted that children with CIs may simply have greater difficulty recognizing degraded speech signals than children with NH.

A. Current study

The primary goal of this study was to examine the abilities of children with CIs to recognize sine-wave speech. These signals provide information about the time-varying structure of the three lowest vocal-tract formants, and this kind of acoustic structure is thought to be especially utilitarian for children in their efforts to discover language structure. Thus, a decrement in sensory availability of time-varying spectral structure can be predicted to have serious deleterious effects on the abilities of these children to recover lexical units, and eventually phonemic units in the ongoing speech signal. Of course, children with CIs would be expected to have poor recognition of any degraded signal. Therefore, any decrement in recognition of sine-wave signals observed for children with CIs, compared to children with NH, could simply reflect poorer abilities to handle, perceptually speaking, degraded signals. Consequently, recognition of sine-wave signals was compared to recognition of the same materials used to create those sine-wave signals, when presented in noise. If it were found that the decrement in performance with sine-wave signals for children with CIs compared to those with NH was similar to that observed between the two groups for speech in noise, it would not indicate specifically that children with CIs had difficulty recovering time-varying formant structure. Rather, it would indicate only that these children are poor at recognition of degraded speech signals.

A second goal of the current study was to examine whether children’s abilities to recognize speech with only sine-wave replicas of the original materials are related to their lexical knowledge or sensitivity to word-internal phonemic structure. Sine-wave signals are of interest because they preserve almost exclusively time-varying formant structure. It has been suggested that children use that structure to parse time-limited sections from the ongoing signal in order to begin constructing early lexicons. Once the rudiments of a lexicon are established, most models of language development suggest that children begin to discover the internal structure of the items in that lexicon, namely, phonemic structure. Given this developmental perspective, the ability to recognize and use time-varying formant structure, as represented with sine-wave replicas of speech, should be related to children’s lexical knowledge and/or phonemic sensitivity (i.e., awareness). In the current study, this was explored using correlation analyses between recognition of sine-wave speech and measures of vocabulary and phonemic awareness.

A third goal of the current study was to examine the abilities of children with CIs to use syntactic and semantic language constraints in their speech recognition, and compare their abilities to those of children with NH. Effects of these “top-down” constraints on speech recognition by children with CIs have been examined before, with outcomes suggesting that children with CIs are poor at applying syntactic and semantic constraints to their recognition of words in sentences (Conway *et al.*, 2014; Eisenberg *et al.*, 2002; Smiljanic and Sladen, 2013). However, those findings have been confounded either by the two groups of children (those

with NH and those with CIs) having greatly different overall recognition probabilities, or by having stimuli that differed in either level of presentation or in quality of the signal across the two groups. These confounds are difficult to avoid when comparing these two groups of children, but nonetheless could influence outcomes. Consequently, it seemed worthwhile to examine top-down language constraints in this study, where signal quality (i.e., sine-wave replicas of speech) was held constant across the two groups of children.

Finally, a fourth goal of this study was to see if factors associated with the children with CIs accounted for variability in their recognition of sine-wave speech. If it were found that sensitivity to this aspect of acoustic speech structure is related to lexical or phonemic knowledge, it would be clinically useful to know how to facilitate children's abilities to recover and use this global spectral structure. Factors examined included age of receiving a first CI, pre-implant auditory thresholds, number of CIs worn, and whether the child had any experience with combined electric-acoustic stimulation.

II. METHODS

A. Participants

Ninety-one children participated in this study: 46 with NH and 45 with severe-to-profound hearing loss who wore CIs. All children had just completed fourth grade at the time of testing, and all were participants in an ongoing longitudinal study involving children with hearing loss (Nittrouer, 2010). At the time of testing, mean age [and standard deviations (SDs)] of the children with NH was 10 years, 5 months (4 months) and mean age (and SDs) of the children with CIs was 10 years, 8 months (5 months). This difference was statistically significant, $t(89) = 2.82$, $p = 0.004$, reflecting the fact that the children with CIs were on average a few months older than the children with NH. Because all children were at the same academic level, this was not considered problematic.

Children were well-matched on socioeconomic status. The metric used to make that assessment was one that has been used before, in which occupational status and highest educational level are ranked on scales from 1 to 8, from lowest to highest, for each parent in the home. These scores are multiplied together, for each parent, and the highest value obtained is used as the socioeconomic metric for the family (Nittrouer and Burton, 2005). According to this scale, means (and SDs) for the children with NH and CIs were 35 (13) and 32 (11), respectively. This difference was not statistically significant. Scores suggest that the average child in the study had at least one parent who had obtained a four-year university degree. None of the children in the study had any disabilities (other than hearing loss) that on their own would be expected to negatively impact language learning.

All children had been given the Leiter International Performance Scales—Revised (Roid and Miller, 2002) two years earlier. This instrument provides a nonverbal assessment of cognitive functioning. All children were found to perform within normal limits on this assessment, with means (and SDs) for the children with NH and CIs of 105 (14) and

99 (18), respectively. This difference was not statistically significant.

Regarding children with CIs, mean age of identification of hearing loss was 6 months (7 months), and mean better-ear pure-tone average (PTA) thresholds for the three frequencies of 0.5, 1.0, and 2.0 kHz before implantation were 105 dB (14 dB) hearing level. Twenty-one of these children had at least one year of experience wearing a hearing aid on the ear contralateral to the ear that received the first CI (i.e., bimodal experience) at the time of receiving that first CI, and 13 of those children eventually received a second CI. In fact, at the time of testing, 27 children wore two CIs. Four children with some bimodal experience stopped wearing a hearing aid before this testing occurred, but did not receive a second CI. Four children with some bimodal experience were still using a hearing aid at the time of testing. Mean age of receiving the first CI was 22 months (18 months), and mean age of receiving the second CI was 49 months (21 months).

B. Equipment

Sentence materials, including the speech-in-noise and sine-wave stimuli, were presented through a computer, with a Creative Labs SoundBlaster soundcard using a 44.1 kHz sampling rate and 16-bit digitization. A Roland MA-12C powered speaker was used, placed one meter in front of the child at zero degrees azimuth. For the phonemic awareness task, stimuli were presented in audio-visual format with the same audio set up as that used for sentences, and a 1500-kbps video signal with 24-bit digitization. All testing was video-audio recorded using a SONY HDR-XR550V video recorder, and children wore Sony FM transmitters to ensure good sound quality on the recordings. Receivers for these FM systems connected to the cameras.

C. Stimuli

Two kinds of sentences were used in this task: (1) four-word sentences that are syntactically correct, but semantically anomalous (e.g., *Dumb shoes will sing*), and (2) five-word sentences that are syntactically correct and semantically informative (e.g., *The book tells a story*). The four-word sentences were originally developed by Boothroyd and Nittrouer (1988) and the five-word sentences were culled from the HINT corpus (Nilsson *et al.*, 1994). Hereafter the four-word sentences are termed *low context* and the five-word sentences are termed *high context*. The high-context sentences included function words, but the low-context sentences did not. Both kinds of sentences have been used extensively with children, so are known to be within children's abilities to recognize (e.g., Eisenberg *et al.*, 2000; Nittrouer and Boothroyd, 1990; Nittrouer *et al.*, 2009; Nittrouer and Lowenstein, 2010). For this study, these sentences were all produced by a male talker with a Midwestern dialect. All sentences used are listed in Appendix A. There were 61 high-context sentences and 51 low-context sentences. One sentence in each condition (always the same one) was used for practice.

All these sentences were processed in each of two ways. For the speech-in-noise condition, the long-term average spectrum of each set of sentences (low- and high-context) was computed and used to shape noise for that set. Each sentence was then embedded in a different stretch of noise at each of two signal-to-noise ratios (SNRs): -3 and 0 dB. All children with CIs were presented with sentences at 0 dB SNR. For children with NH, roughly half were presented with sentences at the -3 dB SNR and roughly half were presented with sentences at 0 dB SNR. This was done to try to make overall recognition across groups of children with NH and CIs as similar as possible, and specific SNRs were selected based on outcomes of earlier studies with children where it was observed that children with CIs required roughly a 3 -dB advantage in SNR to obtain similar recognition scores (e.g., Caldwell and Nittrouer, 2013).

For the sine-wave condition, tracks for the first three formants were extracted using a PRAAT routine written by Darwin (2003). However, parameters (such as number of formants to be extracted) were adjusted on a sentence-by-sentence basis to ensure that extracted formants matched those of the original speech files as closely as possible. This was checked by eye, and formant extraction was repeated when necessary. Smoothing of tracks was performed in PRAAT to remove spurious and erroneous excursions. All sentences (original, speech in noise, and sine wave) were equalized so that root mean square amplitude across them was equivalent.

D. General procedures

Test procedures were approved by the Internal Review Board of the Ohio State University, and informed consent was obtained from parents. All children came to the Ohio State University for testing. Children were tested individually in sessions lasting no more than one hour. Breaks of at least one hour were provided between those data-collection sessions. Data for the study reported here were collected across three test sessions: one presenting the sentence materials, one for the phonemic awareness task, and one for the vocabulary task. Other sorts of data were also collected during the latter two sessions, but are not reported here. All materials were presented at 68 dB sound pressure level.

E. Task-specific procedures and materials

1. Sentence materials

Prior to testing in each condition, the practice sentence was presented in its original form, and with the processing being used in that condition (speech in noise or sine wave). Each child heard half the sentences in each set (low or high context) as speech in noise or as sine waves. For each child, the software randomized the selection of sentences to be played in each processing condition. Also for each child, the order of presentation was randomized, such that processing condition alternated, within blocks of low- or high-context sentences. The original, unprocessed sentences were presented after the processed ones, maintaining the alternating pattern of low- or high-context sentences. Thus, an example

of a test order would be 25 low-context sine-wave sentences, 25 low-context speech-in-noise sentences, 30 high-context sine-wave sentences, 30 high-context speech-in-noise sentences, 50 original low-context sentences, and 60 original high-context sentences. Each child was given a cardboard card with six circles on it, and stamped a circle after each condition. In this way children could keep track of how far along in the testing they were.

Children's responses were video-audio recorded. Later, a graduate student scored all responses on a word-by-word basis. Having the video display greatly helped to clarify what the child was saying. However, it also meant that the scorer was not completely blind with respect to children's hearing status because CIs were usually visible. Whole sentences were also scored as correct or not. In order for the sentence to be correct, each of the four or five words had to be repeated correctly, with no additional words. A second student independently scored ten of these recordings (five each from children with NH and those with CIs), and scores between the two students were compared on a word-by-word basis for each child to obtain a metric of reliability. The dependent measure for this task was always the percentage of words recognized correctly.

2. Phonemic awareness

In this study, children's awareness of (or sensitivity to) word-internal phonemic structure was evaluated using a final consonant choice task. It consisted of 48 trials, which are shown in Appendix B. All words were spoken by the same male, English talker with a Midwestern dialect. In each trial, the child first saw and heard the talker say a target word. The child had to repeat the target word correctly, and was given three attempts to do so. After repeating the target word, the child saw and heard three word choices, and had to say which ended in the same sound as the target word. The trials were presented in the same order for each child. The experimenter in the room with the child entered responses as correct or incorrect, and the software automatically ended testing after six consecutive incorrect responses. Testing was video-audio recorded. A member of the laboratory staff (other than the person who was present at testing) viewed 10 of the recordings and independently scored responses for a metric of reliability.

3. Expressive vocabulary

In this study, expressive vocabulary was selected for use rather than receptive vocabulary because it provides a deeper test of vocabulary knowledge. In a standard receptive vocabulary task, the child hears a word and needs to select the picture of that word from among a set of four words. This permits correct responses for words that a child has not completely mastered, ones that have not reached the level of retention in the terminology of fast mapping (e.g., Walker and McGregor, 2013). In an expressive vocabulary task, children see a picture, and must be able to retrieve the correct word label from their own lexicons. This requires that all retrieved words have reached the level of retention.

TABLE I. Means (and SDs) of percent words correct for all sentences. Numbers of children in each group are shown.

	NH -3 dB		NH 0 dB		CI 0 dB	
	21		25		45	
	M	SD	M	SD	M	SD
Low-context sentences						
Speech in noise	46.6	13.3	67.8	9.2	15.6	10.1
Sine wave	33.1	17.1	29.2	17.0	27.2	12.8
High-context sentences						
Speech in noise	83.2	4.2	95.8	2.5	46.3	21.2
Sine wave	90.4	5.1	89.7	6.6	83.8	18.9

In this study, the Expressive One-Word Picture Vocabulary Test, or EOWPVT (Brownell, 2000) was used. The materials in this task consist of a set of easels that are shown one at a time to elicit labeling responses. In this task, testing is discontinued after six consecutive incorrect responses. Again, testing was video-audio recorded and 10 recordings were reviewed by a second student to obtain a measure of reliability.

III. RESULTS

Across the four sentence conditions (low-context speech in noise, low-context sine waves, high-context speech in noise, and high-context sine waves) word-by-word agreement between the two scorers varied from 91% to 100% for individual children. Mean agreement across the ten children scored by each of two scorers in each condition was 98.5% for low-context speech in noise, 97.3% for low-context sine waves, 98.5% for high-context speech in noise, and 99.2% for high-context sine waves. These values were considered to represent adequate reliability. On the measures of phonemic awareness and expressive vocabulary, the second scorer agreed with all scoring done by the experimenter in the room with the child at the time of testing.

Screening of the data showed that all measures had normal distributions, except for the recognition scores for the high-context sentences. These scores tended to be close to 100% correct, so were somewhat negatively skewed. As a consequence, arcsine transformations were used in analyses

for all word recognition scores, in high- and low-context sentences alike. The alpha level for significance was set at 0.05, but p values are reported when $p < 0.10$. When $p > 0.10$, outcomes are reported simply as not significant.

A. Sentence recognition

Table I shows mean percent correct recognition scores for the low- and high-context sentences for each group, with children with NH divided according to the SNR at which speech-in-noise materials were presented. It is clear from these values that children with CIs performed more poorly than children with NH when all were listening to speech in noise at equivalent SNRs (i.e., 0 dB), as well as when children with NH were listening at a poorer SNR (i.e., -3 dB). However, it appears that children with NH and those with CIs performed similarly when listening to sine-wave speech. To examine these impressions, two separate repeated-measures analyses of variance (ANOVAs) were performed: one involving children with NH who heard the speech-in-noise stimuli at -3 dB and one involving children with NH who heard the speech-in-noise stimuli at 0 dB. Type of sentence (low or high context) and processing condition (speech in noise or sine wave) were the repeated measures. Group was the between-subjects measure.

Table II shows the outcomes of the ANOVAs described above. Looking first at the main effects, sentence context and group were significant in both ANOVAs. The main effect of processing condition was significant when children with NH heard the speech-in-noise materials at -3 dB SNR, but not when they heard those stimuli at 0 dB SNR. This difference across SNRs reflects the fact that children with NH performed significantly worse in the speech in noise condition when SNR was -3 dB than when it was 0 dB, both for low-context sentences, $t(44) = 5.44$, $p < 0.001$, and for high-context sentences, $t(44) = 12.28$, $p < 0.001$. Thus scores in the speech-in-noise condition were lowered overall (i.e., across groups) when SNR was -3 dB, evoking a significant processing effect only at this SNR.

Looking at interactions, the sentence context \times processing condition interaction was significant, regardless of which SNR was presented to children with NH. In both cases, recognition was better overall for the speech-in-noise stimuli

TABLE II. Outcomes of ANOVAs performed on word recognition scores. Analyses were performed separately for each SNR presented to children with NH. Type of sentence context = low or high, processing condition = speech in noise or sine wave, group = NH or CI.

	NH -3 dB (df = 1, 64)			NH 0 dB (df = 1, 68)		
	F	p	η^2	F	p	η^2
Main effects						
Type of sentence context	993.98	<0.001	0.940	1458.44	<0.001	0.955
Processing condition	56.16	<0.001	0.467	NS	NS	—
Group	31.04	<0.001	0.327	78.71	<0.001	0.537
Two-way interactions						
Context \times processing	114.43	<0.001	0.641	200.89	<0.001	0.747
Context \times group	NS	NS	—	NS	NS	—
Processing \times group	71.67	<0.001	0.528	264.98	<0.001	0.796
Three-way interaction						
Context \times processing \times group	NS	NS	—	NS	NS	—

TABLE III. Means and SDs for each processing condition with each sentence type, across children with NH and those with CIs at each SNR.

	SNR = -3 dB		SNR = 0 dB	
	21		70	
	M	SD	M	SD
Low-context sentences				
Speech-in-noise	46.6	13.3	34.2	27.0
Sine wave	33.1	17.1	27.9	14.4
High-context sentences				
Speech-in-noise	83.2	4.2	64.0	29.3
Sine wave	90.4	5.1	85.9	15.8

than for the sine-wave stimuli for low-context sentences, but better overall for the sine-wave than the speech-in-noise stimuli with the high-context sentences. These outcomes are shown in Table III. The interaction of sentence context \times group was not significant, indicating that children with NH and those with CIs showed the same proportional improvement for high-context over low-context sentences, so data in Table III are collapsed across listener groups. This finding (of a lack of a context \times group interaction) is the first piece of evidence suggesting that top-down constraints exerted effects of similar magnitude across groups. However, the most important interaction found in the ANOVAs reported here were the significant processing \times group interactions, found at both SNRs. This interaction reveals that children with NH recognized sentences better in the speech-in-noise condition than in the sine-wave condition, regardless of which SNR they were presented with, whereas children with CIs heard sentences better in the sine-wave condition. To illustrate this interaction, mean scores were computed for each child across the low-context and high-context sentences for each of the speech-in-noise and sine-wave conditions. Because children in both groups showed similar effects of sentence context, computing these cross-sentence means was appropriate, and illustrative. Group means are displayed in Table IV and reveal the trends described above.

It could be argued, of course, that a SNR could be found at which even the children with CIs would show better recognition for the speech-in-noise than for the sine-wave condition, making this comparison across processing conditions irrelevant. However, a major outcome of the current investigation was that children with CIs performed similarly to children with NH when speech materials were equivalently degraded in one manner (i.e., sine waves), but not when they were equivalently degraded in another manner (i.e., speech

TABLE IV. Means and SDs for each group, in each processing condition across both types of sentences (low and high context).

	NH -3 dB		NH 0 dB		CI 0 dB	
	21		25		45	
	M	SD	M	SD	M	SD
Speech in noise	64.9	7.1	81.8	5.2	30.9	15.1
Sine wave	61.8	9.5	59.4	10.0	55.5	14.7

TABLE V. Outcomes of t tests performed on means across sentence types (i.e., low- and high-context), at each SNR presented to children with NH. At each SNR (presented to children with NH), performance between children with NH and children with CIs was compared.

	NH -3 dB (df = 64)		NH 0 dB (df = 68)	
	t	p	t	p
Speech in noise	9.78	<0.001	16.28	<0.001
Sine wave	1.77	0.082	NS	NS

in noise). To illustrate this finding, a series of t tests were performed on the cross-sentence means shown in Table IV. These were done on recognition scores for each SNR at which children with NH heard the speech-in-noise stimuli. Outcomes are shown in Table V, and reveal that regardless of which SNR children with NH were listening at, they performed better than children with CIs for the speech-in-noise sentences. However, there was no group effect found for the sine-wave sentences. This is considered a remarkable outcome because there is no other situation in which children with CIs have been shown to perform similarly to children with NH. In fact, outcomes from this study for the unprocessed materials are shown in Table VI, and children with CIs performed significantly more poorly than children with NH for these materials: for low-context sentences, $t(89) = 11.10$, $p < 0.001$; for high-context sentences, $t(89) = 5.68$, $p < 0.001$.

In summary, these results demonstrate that children with CIs are able to recover time-varying formant patterns as well as children with NH, at least when those formants were represented as sine-wave signals.

B. Relationship of sine-wave perception to phonemic awareness and expressive vocabulary

Another goal of the current investigation was to see if children's abilities to utilize the spectral structure available in the sine-wave signals were related to their lexical knowledge and/or sensitivity to phonemic structure. Table VII shows means for the measures of expressive vocabulary (i.e., EOWPVT standard scores) and phonemic awareness (i.e., percent correct on the final consonant choice task). For these measures, t tests revealed that children with CIs performed more poorly than children with NH: for expressive vocabulary, $t(89) = 3.65$, $p < 0.001$; for phonemic awareness, $t(89) = 5.21$, $p < 0.001$.

To accomplish the goal of examining relationships between recognition of sine-wave signals and these other

TABLE VI. Means and SDs of recognition scores for unprocessed materials, presented in quiet.

	NH		CIs	
	46	45	45	45
	M	SDs	M	SDs
Low-context sentences	94.9	4.6	70.8	17.3
High-context sentences	99.3	0.9	94.7	7.8

TABLE VII. Means and SDs for vocabulary and phonemic awareness measures. Expressive vocabulary is given in standard scores and phonemic awareness is percent correct.

	NH		CIs	
	46		45	
	M	SDs	M	SDs
Expressive vocabulary	107	11	96	17
Phonemic awareness	80	16	57	25

measures, Pearson product-moment correlation coefficients were computed between recognition scores for the sine-wave materials, and standard scores on the expressive vocabulary task and percent correct scores on the phonemic awareness task. Because the context \times group interaction had not been significant, the cross-context individual means summarized in Table IV were used in the analyses.

Table VIII shows these correlation coefficients, for all children together, and for each group separately. These results indicate that phonemic awareness, but not expressive vocabulary, was related to children's abilities to recognize sentences in the sine-wave condition. Figure 2 shows a scatterplot of this significant relationship. It is difficult to ascertain with certainty the direction of relationship between phonemic awareness and recognition of sine-wave sentences, but it seems appropriate to reason that if phonemic awareness accounted strongly for children's recognition of these degraded signals, recognition across the two processing conditions would be similar for children with CIs, but it was not. Instead, children with CIs were disproportionately better at recognizing sine-wave signals than speech in noise. Those abilities indicated how well these children were able to recover and use time-varying formant structure. It appears that those abilities then facilitated children's abilities to discover phonemic structure in the speech signal.

C. Top-down language constraints

A third goal of the current study was to examine the abilities of children with NH and those with CIs to apply their knowledge of top-down language constraints to their speech perception. The current study provided an especially good opportunity to conduct this investigation because the same signals (i.e., sine waves) were used with both children with NH and those with CIs, and children in both groups showed similar recognition scores for these signals. Thus, signal structure and performance were equivalent.

TABLE VIII. Pearson product-moment correlation coefficients between across-context means of percent correct recognition for sine-wave materials and measures of expressive vocabulary and phonemic awareness.

	All	NH	CIs
	91	46	45
Expressive vocabulary	0.132	0.244	-0.012
Phonemic awareness	0.522 ^a	0.366 ^b	0.554 ^a

^a $p < 0.01$.

^b $p < 0.05$.

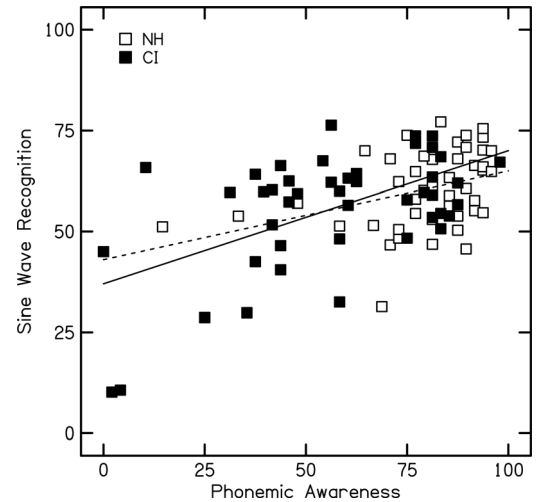


FIG. 2. Scatterplot of relationship between word recognition for sinewave sentences and phonemic awareness. The solid line represents the regression for children with CIs, and the dotted line represents the regression for children with NH.

For this purpose, two metrics were applied to recognition scores for the sine-wave sentences. First, j factors, developed by Boothroyd (e.g., Boothroyd, 1968; Boothroyd and Nittrouer, 1988), were included. With this metric, the number of independent channels of information required to recognize the sentence are computed based on the equation

$$p_s = p_p^j \quad (1)$$

In this formula, p_s is the proportion of whole sentences recognized correctly and p_p is the proportion of parts, or words, recognized correctly. The formula grows out of the idea that in the absence of top-down constraints, every word would need to be recognized correctly in order for the whole sentence to be recognized correctly, and the exponent j would equal the number of words in the sentence. However, because listeners can apply their knowledge of top-down constraints to aid recognition, the number of words that must be recognized correctly decreases, so j becomes less than the number of words in the sentence. It now is viewed more appropriately as indexing the number of independent channels of information—be they lexical, syntactic, or semantic—that are required for sentence recognition. This value can be computed with the following equation, derived from Eq. (1):

$$j = \log(p_s) / \log(p_p) \quad (2)$$

In this study, j was computed separately for the low- and high-context sentences, in the sine-wave condition.

In addition to j , the difference in word recognition scores (in percent correct) between the low- and high-context sentences for the sine-wave stimuli was computed for each child. Although the same words were not used in the construction of sentences in each set, the high-context sentences were semantically rich, so the contributions of this kind of linguistic structure could be evaluated by comparing differences in scores between the two kinds of sentences.

TABLE IX. Means and SDs for metrics of top-down language constraints: j factors for both sentence types, and difference scores between word recognition for low- and high-context sine-wave sentences.

	NH		CIs	
	46		45	
	M	SD	M	SD
Low-context sentences j factor	2.67	0.60	2.55	0.46
High-context sentences j factor	3.32	0.96	3.77	1.80
Difference score	59.0	16.4	56.6	13.1

Table IX shows group means for each of these three metrics of top-down effects. None of them showed a significant group effect, so the conclusion may be reached that children with CIs appeared to use top-down constraints to the same extent as children with NH.

D. Audiological factors contributing to outcomes

The fourth and final goal of this study was to examine whether or not there were any audiological factors that accounted for how well children with CIs could recognize the degraded signals, including both the speech-in-noise and the sine-wave materials. To address this goal, a series of 16 Pearson product-moment correlation coefficients were computed, between recognition scores for the speech-in-noise and sine-wave materials, with both sentence types, and the treatment factors of age of identification, age of first implant, age of second implant (for 27 children), and pre-implant better-ear PTAs. None of these correlation coefficients was significant.

Next, the effects of having one or two CIs at the time of testing, and of having had or not had a period of bimodal experience at the time of receiving a first CI were examined. Table X shows mean recognition scores and SDs for children with CIs, as a function of whether they had one or two CIs and whether they had a year or more of bimodal experience at the time of receiving their first CI. The four children who continued to use bimodal stimulation at the time of testing are not included here because there were too few of them to form a meaningful group on their own and they did not fit neatly into the other groupings. Two-way ANOVAs were performed separately on scores for the four conditions shown

on each row of Table X, with number of CIs and bimodal experience as the between-subjects factors. The only significant effect found involved whether or not children had some bimodal experience, and this effect was significant only for the sine-wave materials: for low-context sentences, $F(1,37) = 4.87$, $p = 0.034$, $\eta^2 = 0.116$; for high-context sentences, $F(1,37) = 6.89$, $p = 0.013$, $\eta^2 = 0.157$. Thus, it can be concluded that children who had a period of bimodal experience were more sensitive to this spectral structure than the children with no bimodal experience. The finding of a lack of effect for recognition of speech-in-noise materials reveals that this effect was not simply due to these children with bimodal experience being generally more skilled perceptually than the children without that experience. Rather, the effect seems to relate specifically to spectral structure; it may be that the enhanced opportunity of the children with some bimodal experience to hear spectral structure as infants helped sensitize them to that structure. This effect was not large, but it could be clinically meaningful. That suggestion is supported by the additional finding that the two groups differed in their scores on the phonemic awareness task: children with some bimodal experience had means (and SDs) of 65.7% correct (22.6% correct) and children without bimodal experience had means (and SDs) of 48.1% correct (25.0% correct). This difference was statistically significant, $t(39) = 2.33$, $p = 0.025$. Given the significant relationship found between recognition of sine-wave materials and phonemic awareness, it is reasonable to assume that this group difference reflects that relationship.

IV. DISCUSSION

The primary purpose of the current study was to examine how well children with CIs could recognize sentences generated from sine-wave replicas of natural speech. The motivation for this investigation was based on the fact that these signals represent especially well the relatively slowly changing patterns of vocal-tract formants. It has been hypothesized that sensitivity to this kind of spectral structure is a prerequisite for children to be able to discover various kinds of linguistic structure, such as words and phonemic units. Going into this study, it was anticipated that children with CIs would have more difficulty than children with NH recognizing sentences generated with sine-wave replicas of

TABLE X. Means and SDs for word recognition in each condition as a function of whether children with CIs had one or two CIs at the time of testing, and whether they had a period of bimodal experience at the time of receiving a first CI. The four children who continued to use bimodal stimulation at the time of testing are not included here.

	One CI		Two CIs		Some bimodal		No bimodal	
	14		27		18		23	
	M	SD	M	SD	M	SD	M	SD
Low-context sentences								
Speech in noise	16.0	9.8	15.6	10.4	17.2	10.5	14.6	9.8
Sine wave	26.6	10.9	28.9	14.0	32.2	9.1	25.0	14.7
High-context sentences								
Speech in noise	45.7	20.3	47.7	22.5	53.2	20.5	42.2	21.4
Sine wave	86.2	13.8	82.5	22.2	92.4	5.2	77.0	23.9

speech. Nonetheless, it was predicted that the abilities of these children to recognize those sine-wave sentences would be strongly correlated with either their lexical knowledge or their sensitivity to phonemic structure. Consequently, tests of lexical knowledge and phonemic awareness were included in this study.

The finding of most significance in this study was simply that children with CIs were able to recognize sine-wave sentences as well as children with NH. This outcome differed from what was found for speech-in-noise, where children with NH performed much better than those with CIs. Nonetheless, the finding of equivalent recognition for sine-wave speech did not come about solely due to children with NH performing worse with sine waves than with speech in noise. Although children with NH did perform somewhat more poorly with sine waves than with speech in noise, the performance of children with CIs for sine waves was much better than their performance for speech in noise. Overall these children with CIs were apparently quite capable of recovering time-varying spectral structure from the sine-wave signals and using that structure to recognize speech.

The second finding of interest in the current study was that children's abilities to recognize the sine-wave signals were significantly related to their phonemic awareness. This relationship was especially strong for the children with CIs, a finding that could reflect the fact that children with NH are able to utilize other kinds of structure in the acoustic speech signal, such as brief, but relatively steady-state spectral patterns; children with CIs cannot use that structure as competently (Nittrouer *et al.*, 2014). Consequently, whatever sensitivity children with CIs have to phonemic structure, it must be based almost exclusively on these relatively slow patterns of spectral change. That finding is a significant outcome of the current study, because it advances our understanding of language acquisition in children with CIs. Recognition scores for sine-wave sentences were not correlated with vocabulary knowledge, for either group of children. That finding could reflect the fact that these children were 10 years old. By this age, it would be expected that their lexicons had become re-organized according to whatever sensitivity to phonemic structure they had. It was, nonetheless, important to include lexical knowledge in the analysis to see if a relationship existed.

The third goal of this study was to see if evidence could be found regarding the abilities of children with CIs to apply their knowledge of top-down language constraints to their sentence recognition, as compared to the abilities of children with NH to do so. The current study provided a fresh opportunity for exploring this question because the signals in the case of sine-wave speech were the same for the two groups of children. As it turned out, recognition scores for children in the two groups were in the same range for these signals, and that further enhanced the validity of the comparison. With these considerations met, it was observed that children with CIs were able to apply their knowledge of top-down constraints to their speech recognition to a similar extent as children with NH.

Finally, a goal of this study was to see if there were any factors related to the hearing loss itself or its treatment that

could account for the abilities of children with CIs to use global spectral structure in their speech perception. This information would be especially useful if it were found—as it were—that sensitivity to this global spectral structure is related to children's vocabulary knowledge or phonemic awareness. That outcome would mean that there would be clinical utility in implementing efforts to facilitate children's abilities to recover this kind of structure. In this study it was observed that children who had some, albeit brief, period of combined electric-acoustic (i.e., bimodal) stimulation were more sensitive to global spectral structure than children who did not have a period of this sort of stimulation. On the other hand, the factors of age of receiving a CI or pre-implant PTAs did not influence sensitivity to global spectral structure. These findings suggest that it could be advantageous to provide a period of bimodal stimulation to children, around the time of their first implantation.

Overall, the outcomes of the current study indicate that the kind of acoustic structure that served as the focus of the current study is facilitative for language acquisition. Consequently, efforts to ensure that this structure is available through CIs for individual children should be useful. These efforts could involve developing processing algorithms that emphasize that kind of structure, as well as diagnostic and mapping procedures that evaluate how well that kind of structure is being represented. New diagnostic tools, such as spectral modulation detection, might help meet this goal (e.g., Gifford *et al.*, 2014).

A. Summary

Lexical and phonemic units do not arrive at the ear in neatly packaged form, so children must discover how to extract them from the acoustic speech signal. This developmental process is protracted, taking place over roughly the first decade of life. One kind of acoustic structure that seems to provide early access to linguistic units involves time-varying patterns of vocal-tract formants, but that kind of structure may not be available to children with CIs. The major goal of this study was to examine whether children with CIs have the same degree of access to this structure as children with NH. To meet that goal, children were asked to recognize sine-wave replicas of sentences, along with the same materials presented in noise. In addition, vocabulary knowledge, phonemic awareness, and use of top-down language constraints were all assessed. Finally, treatment factors were examined as possible predictors of outcomes. Results showed that children with CIs were as accurate as children with NH at recognizing sine-wave speech, but were poorer at recognizing speech in noise. Phonemic awareness was significantly related to that sine-wave recognition, and top-down effects were similar in magnitude across groups. Having had a period of bimodal stimulation near the time of receiving a first CI facilitated all these effects for children with CIs. Results suggest that efforts should be made to ensure that all children with CIs have access to the time-varying patterns of vocal-tract formants.

ACKNOWLEDGMENTS

The authors wish to thank Amanda Caldwell-Tarr, Caitlin Rice, Eric Tarr, Jill Twersky, Taylor Wucinich, and Kovid Bhatnagar for help in data collection. This work was supported by Grant No. R01 DC006237 from the National Institutes of Health, National Institute on Deafness and Other Communication Disorders.

APPENDIX A: SENTENCES USED TO CREATE THE SINE-WAVE AND SPEECH-IN-NOISE MATERIALS

P = practice. High-context sentences.

P. The two farmers were talking

- | | |
|-----------------------------------|--|
| 1. Flowers grow in the garden. | 31. The police helped the driver. |
| 2. She looked in her mirror. | 32. He really scared his sister. |
| 3. They heard a funny noise. | 33. He found his brother hiding. |
| 4. The book tells a story. | 34. She lost her credit card. |
| 5. The team is playing well. | 35. He wore his yellow shirt. |
| 6. The lady packed her bag. | 36. The young people are dancing. |
| 7. They waited for an hour. | 37. Her husband brought some flowers. |
| 8. The silly boy is hiding. | 38. The children washed the plates. |
| 9. The mailman shut the gate. | 39. The baby broke his cup. |
| 10. The dinner plate is hot. | 40. They are coming for dinner. |
| 11. They knocked on the window. | 41. They had a wonderful day. |
| 12. He is sucking his thumb. | 42. The bananas were too ripe. |
| 13. He grew lots of vegetables. | 43. She argues with her sister. |
| 14. He hung up his raincoat. | 44. The kitchen window was clean. |
| 15. The mother heard the baby. | 45. The mailman brought a letter. |
| 16. The apple pie was good. | 46. He climbed up the ladder. |
| 17. New neighbors are moving in. | 47. He is washing his car. |
| 18. The woman cleaned her house. | 48. The sun melted the snow. |
| 19. The old gloves are dirty. | 49. The scissors are very sharp. |
| 20. The painter uses a brush. | 50. Swimmers can hold their breath. |
| 21. The bath water is warm. | 51. The boy is running away. |
| 22. Milk comes in a carton. | 52. The driver started the car. |
| 23. The ball bounced very high. | 53. The children helped their teacher. |
| 24. School got out early today. | 54. The chicken laid some eggs. |
| 25. The rain came pouring down. | 55. The ball broke the window. |
| 26. The train is moving fast. | 56. Snow falls in the winter. |
| 27. The baby slept all night. | 57. The baby wants his bottle. |
| 28. Someone is crossing the road. | 58. The orange is very sweet. |
| 29. The big fish got away. | 59. The oven door was open. |
| 30. The man called the police. | 60. The family bought a house. |

Low-context sentences

P. Cooks run in brooms

- | | |
|-----------------------------|--------------------------------|
| 1. Hot slugs pick boats. | 14. Feet catch bright thieves. |
| 2. Wide pens swim high. | 15. Cats get bad ground. |
| 3. Dumb shoes will sing. | 16. Sad cars want chills. |
| 4. True kings keep new. | 17. Leave them cool fun. |
| 5. Blocks can't run sharp. | 18. Hard corn feels mean. |
| 6. Drive my throat late. | 19. Knees talk with mice. |
| 7. Drums pour tall pets. | 20. Late forks hit low. |
| 8. Stars find clean roof. | 21. Lend them less sleep. |
| 9. Tame beans test ice. | 22. Paint your belt warm. |
| 10. Green hands don't sink. | 23. Big apes grab sun. |
| 11. Bad dogs sail up. | 24. Teeth sleep on doors. |
| 12. Socks pack out ropes. | 25. Small lunch wipes sand. |
| 13. Suits burn fair trail. | 26. Late fruit spins lakes. |

(Continued)

- | | |
|-----------------------------|------------------------------|
| 27. Hard checks think tall. | 39. Blue chairs speak well. |
| 28. Tin hats may laugh. | 40. Slow dice buy long. |
| 29. Soap takes on dogs. | 41. Lead this coat home. |
| 30. Cars jump from fish. | 42. Pink chalk bakes phones. |
| 31. They turn small trees. | 43. Shy laws have keys. |
| 32. Trucks drop sweet dust. | 44. High bears move holes. |
| 33. Let their flood hear. | 45. Call her wing guide. |
| 34. Long kids stay back. | 46. Four rats kick warm. |
| 35. Guys tell loud meat. | 47. Soft rocks taste red. |
| 36. Thin books look soft. | 48. Cold worms have toys. |
| 37. Snow smells more tough. | 49. Fan spells large toy. |
| 38. Cups kill fat leaves. | 50. Jobs get thick hay. |

APPENDIX B: THE FINAL CONSONANT CHOICE TASK USED TO ASSESS PHONEMIC AWARENESS

Practice items

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

Discontinue after six consecutive errors

Test trials

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

Beckman, M. E., and Edwards, J. (2000). "The ontogeny of phonological categories and the primacy of lexical learning in linguistic development," *Child Dev.* **71**, 240–249.

Boothroyd, A. (1968). "Statistical theory of the speech discrimination score," *J. Acoust. Soc. Am.* **43**, 362–367.

Boothroyd, A., and Nittrouer, S. (1988). "Mathematical treatment of context effects in phoneme and word recognition," *J. Acoust. Soc. Am.* **84**, 101–114.

- Brownell, R. (2000). *Expressive One-Word Picture Vocabulary Test (EOWPVT)*, 3rd ed. (Academic Therapy Publications, Novato, CA).
- Caldwell, A., and Nittrouer, S. (2013). "Speech perception in noise by children with cochlear implants," *J. Speech Lang Hear. Res.* **56**, 13–30.
- Charles-Luce, J., and Luce, P. A. (1990). "Similarity neighbourhoods of words in young children's lexicons," *J. Child Lang* **17**, 205–215.
- Conway, C. M., Deocampo, J., Walk, A. M., Anaya, E. M., and Pisoni, D. B. (2014). "Deaf children with cochlear implants do not appear to use sentence context to help recognize spoken words," *J. Speech Lang Hear. Res.* **57**, 2174–2190.
- Darwin, C. (2003). Sine-wave speech produced automatically using a script for the PRAAT program. http://www.lifesci.sussex.ac.uk/home/Chris_Darwin/SWS/ (Last viewed October 15, 2014).
- Eisenberg, L. S., Martinez, A. S., Holowecky, S. R., and Pogorelsky, S. (2002). "Recognition of lexically controlled words and sentences by children with normal hearing and children with cochlear implants," *Ear Hear.* **23**, 450–462.
- Eisenberg, L. S., Shannon, R. V., Schaefer Martinez, A., Wygonski, J., and Boothroyd, A. (2000). "Speech recognition with reduced spectral cues as a function of age," *J. Acoust. Soc. Am.* **107**, 2704–2710.
- Ferguson, C. A., and Farwell, C. B. (1975). "Words and sounds in early language acquisition," *Language* **51**, 419–439.
- Gifford, R. H., Hedley-Williams, A., and Spahr, A. J. (2014). "Clinical assessment of spectral modulation detection for adult cochlear implant recipients: A non-language based measure of performance outcomes," *Int. J. Audiol.* **53**, 159–164.
- Jusczyk, P. W. (1995). "Language acquisition: Speech sounds and the beginning of phonology," in *Speech, Language, and Communication*, edited by J. L. Miller and P. D. Eimas (Academic Press, San Diego, CA), pp. 263–301.
- Kuhl, P. K. (1987). "Perception of speech and sound in early infancy," in *From Perception to Cognition, Vol. 2 of Handbook of Infant Perception*, edited by P. Salapatek and L. Cohen (Academic Press, New York), pp. 275–382.
- Lieberman, A. M. (1996). *Speech: A Special Code* (MIT Press, Cambridge, MA).
- Licklider, J. C., and Pollack, I. (1948). "Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech," *J. Acoust. Soc. Am.* **20**, 42–51.
- Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise," *J. Acoust. Soc. Am.* **95**, 1085–1099.
- Nittrouer, S. (2006). "Children hear the forest," *J. Acoust. Soc. Am.* **120**, 1799–1802.
- Nittrouer, S. (2010). *Early Development of Children with Hearing Loss* (Plural Publishing, San Diego).
- Nittrouer, S., and Boothroyd, A. (1990). "Context effects in phoneme and word recognition by young children and older adults," *J. Acoust. Soc. Am.* **87**, 2705–2715.
- Nittrouer, S., and 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," *J. Commun. Disord.* **38**, 29–63.
- Nittrouer, S., Caldwell-Tarr, A., Moberly, A. C., and Lowenstein, J. H. (2014). "Perceptual weighting strategies of children with cochlear implants and normal hearing," *J. Commun. Disord.* **52**, 111–133.
- Nittrouer, S., Caldwell-Tarr, A., Tarr, E., Lowenstein, J. H., Rice, C., and Moberly, A. C. (2013). "Improving speech-in-noise recognition for children with hearing loss: Potential effects of language abilities, binaural summation, and head shadow," *Int. J. Audiol.* **52**, 513–525.
- Nittrouer, S., and Lowenstein, J. H. (2010). "Learning to perceptually organize speech signals in native fashion," *J. Acoust. Soc. Am.* **127**, 1624–1635.
- Nittrouer, S., Lowenstein, J. H., and Packer, R. (2009). "Children discover the spectral skeletons in their native language before the amplitude envelopes," *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 1245–1253.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., and Carrell, T. D. (1981). "Speech perception without traditional speech cues," *Science* **212**, 947–949.
- Roid, G. H., and Miller, L. J. (2002). *Leiter International Performance Scale—Revised (Leiter-R)* (Stoelting Co., Wood Dale, IL).
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Smiljanic, R., and Sladen, D. (2013). "Acoustic and semantic enhancements for children with cochlear implants," *J. Speech Lang Hear. Res.* **56**, 1085–1096.
- Storkel, H. L. (2002). "Restructuring of similarity neighbourhoods in the developing mental lexicon," *J. Child Lang.* **29**, 251–274.
- Treiman, R., and Breaux, A. M. (1982). "Common phoneme and overall similarity relations among spoken syllables: Their use by children and adults," *J. Psycholinguist. Res.* **11**, 569–598.
- Walker, E. A., and McGregor, K. K. (2013). "Word learning processes in children with cochlear implants," *J. Speech Lang Hear. Res.* **56**, 375–387.
- Walley, A. C., Metsala, J. L., and Garlock, V. M. (2003). "Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability," *Read. Writ.* **16**, 5–20.
- Walley, A. C., Smith, L. B., and Jusczyk, P. W. (1986). "The role of phonemes and syllables in the perceived similarity of speech sounds for children," *Mem. Cognit.* **14**, 220–229.
- Werker, J. F. (1991). "The ontogeny of speech perception," in *Modularity and the Motor Theory of Speech Perception*, edited by I. G. Mattingly and M. Studdert-Kennedy (Lawrence Erlbaum Associates, Hillsdale, NJ), pp. 91–109.