

Toddlers' fast-mapping from noise-vocoded speech

Rochelle S. Newman, Giovanna Morini, Emily Shroads, and Monita Chatterjee

Citation: *The Journal of the Acoustical Society of America* **147**, 2432 (2020); doi: 10.1121/10.0001129

View online: <https://doi.org/10.1121/10.0001129>

View Table of Contents: <https://asa.scitation.org/toc/jas/147/4>

Published by the *Acoustical Society of America*

ARTICLES YOU MAY BE INTERESTED IN

[Acoustic factors affecting interaural level differences for cochlear-implant users](#)

The Journal of the Acoustical Society of America **147**, EL357 (2020); <https://doi.org/10.1121/10.0001088>

[Talker variability in word recognition under cochlear implant simulation: Does talker gender matter?](#)

The Journal of the Acoustical Society of America **147**, EL370 (2020); <https://doi.org/10.1121/10.0001097>

[Influence of bilateral cochlear implants on vocal control](#)

The Journal of the Acoustical Society of America **147**, 2423 (2020); <https://doi.org/10.1121/10.0001099>

[Cognitive mechanisms underpinning successful perception of different speech distortions](#)

The Journal of the Acoustical Society of America **147**, 2728 (2020); <https://doi.org/10.1121/10.0001160>

[Mandarin tone perception in multiple-talker babbles and speech-shaped noise](#)

The Journal of the Acoustical Society of America **147**, EL307 (2020); <https://doi.org/10.1121/10.0001002>

[Production-perception relationship of Mandarin tones as revealed by critical perceptual cues](#)

The Journal of the Acoustical Society of America **147**, EL301 (2020); <https://doi.org/10.1121/10.0000963>

CALL FOR PAPERS

JASA
THE JOURNAL OF THE
ACOUSTICAL SOCIETY OF AMERICA

**Special Issue: Ocean Acoustics
in the Changing Arctic**

Toddlers' fast-mapping from noise-vocoded speech

Rochelle S. Newman,^{1,a)} Giovanna Morini,^{2,b)} Emily Shroads,¹ and Monita Chatterjee^{3,c)}

¹Department of Hearing and Speech Sciences, University of Maryland, 0100 Lefrak Hall, College Park, Maryland 20742, USA

²Department of Communication Sciences and Disorders, University of Delaware, 100 Discovery Boulevard, Newark, Delaware 19713, USA

³Boys Town National Research Hospital, 555 North 30th Street, Omaha, Nebraska 68131, USA

ABSTRACT:

The ability to recognize speech that is degraded spectrally is a critical skill for successfully using a cochlear implant (CI). Previous research has shown that toddlers with normal hearing can successfully recognize noise-vocoded words as long as the signal contains at least eight spectral channels [Newman and Chatterjee. (2013). *J. Acoust. Soc. Am.* **133**(1), 483–494; Newman, Chatterjee, Morini, and Remez. (2015). *J. Acoust. Soc. Am.* **138**(3), EL311–EL317], although they have difficulty with signals that only contain four channels of information. Young children with CIs not only need to match a degraded speech signal to a stored representation (word recognition), but they also need to create new representations (word learning), a task that is likely to be more cognitively demanding. Normal-hearing toddlers aged 34 months were tested on their ability to initially learn (fast-map) new words in noise-vocoded stimuli. While children were successful at fast-mapping new words from 16-channel noise-vocoded stimuli, they failed to do so from 8-channel noise-vocoded speech. The level of degradation imposed by 8-channel vocoding appears sufficient to disrupt fast-mapping in young children. Recent results indicate that only CI patients with high spectral resolution can benefit from more than eight active electrodes. This suggests that for many children with CIs, reduced spectral resolution may limit their acquisition of novel words. © 2020 Acoustical Society of America.

<https://doi.org/10.1121/10.0001129>

(Received 30 October 2019; revised 2 April 2020; accepted 3 April 2020; published online 21 April 2020)

[Editor: Deniz Baskent]

Pages: 2432–2441

I. INTRODUCTION

Cochlear implants (CIs) are currently the standard of care for young children born with profound hearing loss. They provide auditory perception by bypassing the damaged portion of the inner ear (the cochlea) and directly stimulating the auditory nerve. Over 38 000 children have been implanted with CIs in the U.S. alone (National Institute on Deafness and Other Communication Disorders, 2016). However, CIs are unable to provide listeners with the complete speech signal. The signal they provide is substantially degraded as a result of both biological and technological limitations (e.g., the loss of auditory neurons from lack of stimulation; cross-talk across neighboring electrodes in the implant; etc.).

Many children with a CI are quite successful at using their device and demonstrate high performance listening to speech in a laboratory setting. Yet, others are far less successful, and researchers have struggled to explain the cause of these differences. In general, children with CIs show substantial variability in both their clinical outcomes and their laboratory performance (e.g., Boons *et al.*, 2012), likely due, in part, to other differences that are hard to control: the exact nature of the hearing loss itself, the child's age at

implantation, atrophy of auditory nerves and changes to the central auditory system from lack of use, parental support, etc. This, combined with difficulty recruiting the population (and the resulting small sample sizes, which exacerbate the impact of intersubject variation), has led some researchers to examine performance of normal-hearing (NH) children listening to *simulated* CI speech (Eisenberg *et al.*, 2000). This approach is quite common in the adult CI literature (e.g., Baskent and Shannon, 2003, 2007; Davis *et al.*, 2005; Friesen *et al.*, 2001; Fu and Shannon, 1999; Fu *et al.*, 1998; Hervais-Adelman *et al.*, 2008; Ihlefeld *et al.*, 2010; Shannon *et al.*, 1995; Shannon *et al.*, 1998; Sheldon *et al.*, 2008). Such research eliminates many of these causes of intersubject variation, allowing an examination of what a listener *could* perceive from a degraded signal if there were no other intervening factors. Such research also serves as a starting point for understanding the challenges faced by children with CIs: if children with typical hearing (who have presumably had longer exposure to sound and spoken language and more opportunities to develop strong learning skills) need a minimum amount of information to be present in the signal in order to interpret it, children with CIs are likely to need at least as much, if not more.

One type of degraded signal frequently used in these simulations is noise-vocoded speech. This signal is thought to be similar in many respects to the signal reaching a CI user. The incoming speech signal is divided into a number

^{a)}Electronic mail: rnewman1@umd.edu, ORCID: 0000-0002-1626-4241.

^{b)}ORCID: 0000-0002-0034-8213.

^{c)}ORCID: 0000-0002-8286-2557.

of distinct, broad frequency bands, and the amplitude of each band is used to modulate a band of noise that covers the same frequency region. The bands are then recombined, resulting in a highly unnatural signal that is nonetheless interpretable as speech (Shannon *et al.*, 1995). Importantly, when the signal is divided into a greater number of bands, more of the spectral resolution in the original signal is preserved. Adult NH listeners can accurately interpret noise-vocoded speech made up of as few as three or four frequency bands (Shannon *et al.*, 1995).

The performance of NH individuals listening to noise-vocoded speech is often used as a comparison to the performance of CI users, although the two populations come to the listening task with very different sets of experiences listening to degraded speech. NH individuals who listen to vocoded speech in an experiment have had a lifetime of listening to full-spectrum speech, and thus their lexical representations are based on a non-degraded signal. In contrast, individuals listening to a CI may not have had much experience listening to full-spectrum (non-degraded) speech, depending on when they were implanted. Those whose hearing loss developed early in life may have lexical representations that are less well-formed, impacting their ability to understand speech. Even adults who were implanted later in life are likely not to have had any recent experience with full-spectrum speech and may frequently hear a more ambiguous speech signal, which could also impact their comprehension. Alternatively, CI users are likely to be able to exploit the degraded signal in ways that those with less experience listening to signal degradation are not. These differences can be taken to make two very different predictions about how NH listeners' performance listening to noise-vocoded speech relates to CI users' performance listening to full-spectrum speech. On one hand, stronger lexical representations that are formed from greater exposure to a non-degraded signal could theoretically result in better performance by individuals listening to a vocoded signal than by individuals with a CI. On the other hand, listeners who have CIs have had (likely many) years of experience listening to this type of degraded signal, possibly allowing them to outperform NH individuals who are listening to vocoded stimuli for the first time. But in fact, when compared directly, the performance of NH adults listening to noise-vocoded speech has proven strikingly similar to the performance of adults with a CI listening to normal (full-spectrum) speech (e.g., Fu and Shannon, 1998, 1999). That is, despite using potentially different mechanisms, the two groups have surprisingly similar outcomes. As a result, much of what we have learned about how CI listeners are able to understand speech through their implant has actually originated from studies using NH listeners with a noise-vocoded signal (see, for example, Baskent and Shannon, 2003, 2006, 2007; Fu and Shannon, 1999; Fu *et al.*, 1998; Hervais-Adelman *et al.*, 2008; Ihlefeld *et al.*, 2010; Sheldon *et al.*, 2008, for work with adults; and Eisenberg *et al.*, 2000, for work with children).

Although most of this work has been done with adult listeners, in recent years we have seen an increase in studies

testing children (Chatterjee *et al.*, 2015; Eisenberg *et al.*, 2000; Newman and Chatterjee, 2013; Newman *et al.*, 2015; Nittrouer and Lowenstein, 2010; Nittrouer *et al.*, 2009). For example, Eisenberg *et al.* (2000) found that children reached adult-like levels of speech recognition (for sentences, phonetically balanced words, and phonetic change detection) by 10–12 years of age, but that children aged 5–7 years required that the signal contain more channels (or frequency bands) in order to recognize it. Newman and Chatterjee (2013) reported that NH toddlers could successfully comprehend noise-vocoded speech as long as the signal contained at least eight channels, but many children failed to recognize known words when the number of channels was down to four channels (a level at which adults are highly successful). This difference based on the number of channels is particularly important because evidence suggests that adult CI listeners are generally not able to benefit from the full number of channels putatively in their device (Fishman *et al.*, 1997). Some data suggest adult CI users may be limited to eight channels of spectral information (Friesen *et al.*, 2001), and some individuals may be receiving as few as four channels. Although this work has been tested with adults, the presumption is that the same would hold true for children. Initial data (Dorman *et al.*, 2000) suggested that school-aged children with CIs perceive speech at a level equivalent to normally hearing adults listening to speech processed through 4–6 spectral channels, again suggesting a limit on the actual number of channels they can benefit from. However, compared to later-implanted children with CIs, earlier-implanted children identified speech at a level equivalent to normally hearing peers' performance with higher numbers of spectral channels (Dorman *et al.*, 2000). Eisenberg *et al.* (2000) showed that younger normally hearing children need more channels of information to achieve the same level of speech recognition as adults. A more recent study (Jahn *et al.*, 2019) on vowel identification by school-aged children with normal hearing and with CIs showed that children with normal hearing are more sensitive to the slopes of the simulated channels (i.e., the degree of channel-interaction) than are adults with normal hearing. Consistent with Dorman *et al.* (2000), they also showed an advantage for earlier-implanted children than for later-implanted children, even though their participants would have been implanted with more modern devices. However, these questions about the number of channels from which children can benefit remain unanswered in very young children.

More recent findings in adult CI users suggest that these limitations may depend on the specifics of electrode placement as well as the specifics of the technology. Individuals with more modern implants in which the electrodes are placed more closely to the spiral ganglion cells demonstrate a shallower asymptote than the earlier studies. Consistent with earlier studies, patients implanted with modern perimodiolar electrode arrays show a large increase in performance from four to eight channels, but with a continuing, smaller yet significant benefit from increasing the number of active

electrodes beyond eight that was not reported previously (Berg *et al.*, 2019; Croghan *et al.*, 2017). This improvement beyond eight channels was demonstrated specifically in patients with smaller electrode-to-neuron distance and better spectral (Berg *et al.*, 2019) or spectro-temporal resolution (Croghan *et al.*, 2017). Consistent with the idea that patients using modern-day devices benefit from increasing the number of channels, Schwartz-Leyzac *et al.* (2017) showed that decreasing the number of active electrodes resulted in a decrease in speech recognition in CI patients. For many CI patients, however, performance with eight electrodes is close to their maximum benefit from the device. Berg *et al.* (2019) further showed that performance with 8 electrodes predicted performance with 16 electrodes active. As the majority of listeners with CIs still gain limited benefit beyond eight channels of speech information, determining how well children can fast-map with this number of channels seems relevant. If young children have difficulty recognizing speech with fewer than eight channels, this may have important limitations on children's success with an implant.

One important caveat is that this prior work examined recognition of words the children were expected to have already learned. Yet, young children with a CI not only need to recognize words but also to learn them initially, a task that is thought to be more cognitively demanding than word recognition (Bloom, 2000) and is potentially impacted to a greater extent by a degraded speech signal. Both recognizing and learning new words require that the sound pattern of the word be identified, but word learning also involves mapping that sound pattern onto an appropriate referent, a task that requires additional working memory skills beyond those required for recognition (Werker and Curtin, 2005). The need to link sound-based and conceptual representations may make the process of learning new words more demanding computationally. Interpreting a degraded signal may likewise require more memory and attentional resources than does listening to a full-spectrum signal (e.g., Mattys *et al.*, 2012; Zekveld and Kramer, 2014), reducing the amount of such resources available for storing information in long-term memory. Because coping with a degraded signal places demands on children's cognitive resources, and word learning depends more heavily on these resources than does word recognition, word learning may be more susceptible to the effects of degradation. That is, we might expect that a degraded signal (such as that from a CI or its analogous vocoded simulation) could lead to poorer learning even when these conditions allow for adequate recognition.

Word learning outside of the laboratory often takes place gradually, over many instances as the child engages in activities of daily life (Bion *et al.*, 2013; Carey, 2010; Kucker *et al.*, 2015; Swingley, 2010). This type of slow learning is difficult to manage in a laboratory setting in a single visit study, although it has been done with multi-visit longitudinal studies. Instead, we assess children's ability to map a label onto an object during a short testing session (or to initially discover and store the appropriate referent of a new word). This word-to-object mapping is an important part of true word learning but likely does not represent the

richness of fully incorporating a word into a semantic network and truly "learning" the word (see, for example, Horst and Samuelson, 2008). However, fast-mapping does have the advantage that it can be studied in a single-visit in a laboratory setting, and the ability appears to relate to "full" word learning (McMurray *et al.*, 2012). Moreover, this initial mapping ability is likely to depend critically on the quality of the signal itself, thus serving as an excellent test case for the impact of signal degradation on word learning.

Children with CIs have been shown to have difficulties with fast-mapping compared to their age-matched peers (Tomblin *et al.*, 2007; Walker and McGregor, 2013). This is particularly the case for children who were implanted at a later age (Houston *et al.*, 2012; Tomblin *et al.*, 2007). Although a number of factors likely contribute to this difficulty (including prior learning as indicated by vocabulary size; Walker and McGregor, 2013), signal degradation is one potential factor.

Thus, in the present study, we test children with normal hearing and normal previous language experience on a fast-mapping task with noise-vocoded stimuli. Specifically, we investigate how well children can initially learn new words from a degraded signal and how much information is needed in the signal for children to be successful. We first train children on the mapping between objects and their word forms using a degraded signal by presenting images of objects individually and repeatedly naming the objects. We then test children by presenting images of those same two objects simultaneously and telling the children to find one of them. We infer that children have learned the mapping if they spend more time looking to the named object compared to the unnamed object. We predict that children will have difficulty fast-mapping new words from a vocoded signal. More specifically, we predict that they will show more difficulty in this fast-mapping task than they have previously shown in word recognition tasks using the same level of signal degradation. (Newman and Chatterjee, 2013).

II. EXPERIMENT 1

This experiment explored whether toddlers could fast-map new words from noise-vocoded speech. We began by testing children listening to eight-channel noise vocoded stimuli since this is the number of channels at which children aged 27 months were highly successful at speech recognition (Newman and Chatterjee, 2013). We taught children names for two novel objects and then tested them on their learning of those word-object mappings. We presumed that if children looked significantly longer at the appropriate image when they were told to look at it vs when they were not, this would be an indication that they had been able to fast-map that word during the training stage despite the reduced signal quality of the speech.

A. Method

1. Participants

Twenty-four children (12 female, 12 male) approximately 34 months of age (range 33 months, 5 days to

34 months, 29 days) participated in this study. There is a rapid increase in the rate of lexical acquisition during the toddler and preschool years (Fenson *et al.*, 1994; McMurray, 2007), making this a particularly relevant age for testing different aspects of word learning. Moreover, this is the same age as the children in a prior fast-mapping study using an identical methodology but full-spectrum speech (Dombroski and Newman, 2014) and slightly older than the 27-month-old children tested on recognition of noise-vocoded stimuli in Newman and Chatterjee (2013). This allows us to be fairly confident that children of this age are capable of both understanding previously known words from a degraded signal, as well as fast-mapping new words from a non-degraded signal. Whether they can successfully fast-map from the degraded signal remains to be seen.

An additional ten children participated but their data were excluded for excessive fussiness/exiting the test area/not attending ($n = 6$), equipment failure ($n = 3$), or experimenter error/failure to record the session ($n = 1$). The children were assigned evenly to one of six stimulus orders (see Sec. II A 3, Procedure). An additional two participants were recruited in the event that additional data would be required in one of the stimulus orders, but their data (the last data collected in these orders) were not ultimately needed. Parents reported that their children had normal hearing and were not currently experiencing symptoms indicative of an ear infection.

The final set of children were 67% Caucasian, 26% African American, 4% Asian, and 4% mixed race. Maternal education averaged 17.1 years: 14 mothers had a master's degree, 9 mothers had a 4-year college degree, and 1 mother had some college. Thus, the participants are from a fairly well-educated, high socioeconomic status background. Parent-reported vocabulary on the Language Development Survey (Rescorla, 1989) ranged from 13 to 305 words. According to parental reports, three children were exposed to another language in the home: one heard 20% Spanish, one heard 5%–10% Spanish, and one heard 10% Portuguese and 10% Spanish.

2. Materials

We taught children two novel words, “coopa” and “needoke,” using vocoded speech in a split-screen preferential looking paradigm (Golinkoff *et al.*, 1987; Hollich, 2006), by means of an identical method to that previously utilized by Dombroski and Newman (2014) to assess the impact of noise on fast-mapping. These word forms were selected because they are both easily discriminable (having highly distinct vowels, consonants that differ in voicing, and different syllable structures) and multisyllabic. We chose to use very discriminable, relatively long word forms to make the task easier for the child, but admittedly this runs the risk of making the laboratory task less realistic in that the children might not need to form as detailed lexical representations in order to succeed as for other words.

The visual stimuli for this study were from Hollich (2006) and consisted of a spiky object on a pedestal and a green multi-limbed creature. Both objects looked like two-dimensional geometric images (as compared to photos) and rotated in three-dimensional space. Assignment of name to object was counterbalanced across participants.

The original audio stimuli were taken from Dombroski and Newman (2014). Noise vocoding was then performed using methods akin to published standards (Shannon *et al.*, 1995). The analysis input range was 200–7000 Hz with a 24 dB/octave rolloff. The signal was then split into eight frequency bands (or channels) using bandpass filtering (Butterworth filters, 24 dB/octave rolloff), and the envelope of each band was extracted using half-wave rectification and low-pass filtering (400 Hz cutoff frequency so children would have reasonable access to F_0 information within the temporal envelope). The envelope derived from each band was then used to amplitude-modulate a white noise signal with the same bandwidth as the original signal band, and these modulated noises were combined at equal amplitude ratios.

3. Procedure

We use the preferential looking paradigm, which measures children's knowledge based on the percentage time they spent looking at a named object. This paradigm has been shown to be both valid and sensitive and is less sensitive to compliance than are standard pointing or speaking tasks (a concern for children in their “terrible twos”; see Golinkoff *et al.*, 2013). Children sat on their caregiver's lap facing a widescreen television (TV). During the training phase, children saw a single object appear on the screen and heard that object being labeled (e.g., “It's a coopa!”). The two objects were labeled in alternation for eight trials (four trials per object). This was followed by a single silent trial, in which both objects appeared together, intended to introduce the idea that objects would now occur on the left and right sides of the screen.

Following this was an eight-trial test phase designed to assess children's learning of those word-object mappings. On these trials, both objects appeared on the screen at the same time, and the speaker instructed the child to look at one of the two objects (“Find the coopa!”). These test stimuli were vocoded in the same manner as the training stimuli (so both familiarization and test items were vocoded, analogous to the fact that both learning and later recognition would be degraded by a CI). Participants were assigned to one of four different trial orders, which counterbalanced which visual object was referred to by which name and which appeared on the left (vs right) of the screen, and which had different randomizations of the eight test trials.

4. Coding

A digital camera recorded each child's eye gaze throughout the study at a rate of 30 frames per second. Two assistants, blind to trial type, individually coded each child's looking behaviors offline on a frame-by-frame basis using

Supercoder coding software (Hollich, 2005). From this, the infants' total duration of looking at each of the two images on each trial was calculated. The first 16 frames (480 ms) occurred before the onset of the target word and were thus ignored. On any trial in which the coders disagreed by more than 15 frames (0.5 s), a third coder was added; this occurred on 47 of the 408 trials. When this happened, the averages of the two closest codings were used as the final data. From this, we determined the percentage of time the child spent looking at the appropriate (named) object on each trial (e.g., looking at the coopa when the coopa was named, and looking at the needoke when the needoke was named) starting from the onset of the first repetition of the target word (Dombroski and Newman, 2014).

B. Results and discussion

We began by ensuring that our items did not generate any particular biases that would be likely to impact the results. During the training trials, children attended equally long to trials containing each of the two objects [spiky object = 89.8% of the possible looking time, multi-limbed object = 88.5%; $t(23) = 0.59, p = 0.56$] and equally long when the item they were viewing was labeled as the *coopa* (87.7%) vs the *needoke* [90.6%; $t(23) = 1.27, p = 0.22$]. During the baseline trial, in which both objects were presented without labeling, they likewise attended for similar amounts of time to each of the two objects [the item that previously had been labeled *coopa* vs *needoke*, 47.0% vs 53.1%, $t(23) = 1.38, p = 0.18$; spiky vs multi-limbed, 47.7% vs 52.5%, $t(23) = 1.08, p = 0.29$]. Thus, the two objects and names appeared relatively well-matched.

We then proceeded to examine the results from the test trials (see the left side of Fig. 1). Children looked to the correct object 53.05% of the time. This was not significantly different from chance [50%; $t(23) = 1.56, p = 0.13$]. Both of the two objects showed a similar pattern [*coopa*, 53.8% accuracy, $t(23) = 1.12, p = 0.27$; *needoke*, 52.3% accuracy, $t = 0.80, p = 0.43$]. Whereas prior research suggests that children of this age can fast-map new words with full-spectrum speech, both in quiet and in noise (Dombroski and Newman, 2014), and can (on average) recognize already-known words that have been degraded through noise-vocoding (Newman and Chatterjee, 2013), the children tested here appear to have significant difficulties with the initial stage of word learning when faced with a degraded signal. Given the variability in performance, it is possible that some children were, in fact, able to learn the words despite the degradation. However, as a group, the children seem to find this a difficult task.

This eight-channel noise-vocoded signal is thought to be a good representation of the quality of the signal heard by most CI recipients: while modern CIs offer putatively more than eight channels of stimulation, cross-channel interference and loss of auditory neurons limits the number of separate channels that listeners can utilize, such that the average individual's performance saturates or begins to

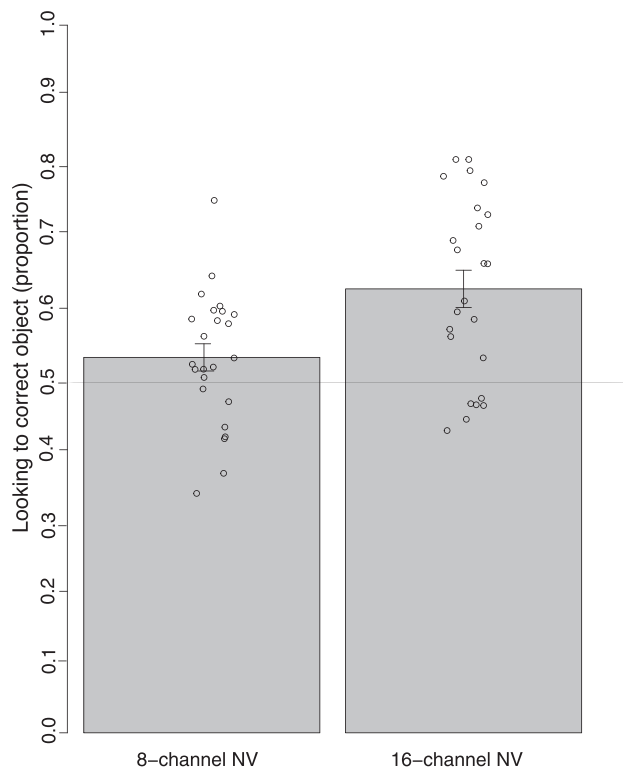


FIG. 1. Proportion looking to the correct object for children in 8-channel noise-vocoded speech (experiment 1) and 16-channel noise-vocoded speech (experiment 2); the line represents chance performance, and circles represent data from individual children.

saturate at about eight channels (Friesen et al., 2001). Thus, the fact that children in the present study did not fast-map words with this type of signal is of concern. That said, it is important to note that the current task examined only fast-mapping not the slow, gradual learning that typically occurs outside of the laboratory setting; difficulty fast-mapping could slow down learning but not necessarily prevent it.

Another possibility, however, is that there was something unusual about this task that prevented children from being successful. This seems unlikely because children of this same age perform well with full-spectrum speech both presented in quiet and in noise (Dombroski and Newman, 2014). While we did not test our current participants with a clear speech condition, our methods for testing and general properties of our population for recruitment are identical to those in Dombroski and Newman (2014), so there is no reason to expect that these older children would have any difficulty with clear speech. However, to ensure that the presence of degradation, in general, was not to blame for children's failure to fast-map new words, we tested another group of children on a less-degraded version of the same stimuli. An additional purpose of this second study was to investigate whether improved spectral resolution might improve learning. Strong intersubject variability is observed in children with CIs, and the results obtained with less severe degradation might be relevant to children listening with less peripheral channel-interaction (e.g., due to a better electrode–neuron interface or better neural survival).

III. EXPERIMENT 2

This experiment is identical to that of experiment 1, except that we presented children with 16-channel noise-vocoded speech rather than 8-channel noise-vocoded speech.

A. Method

1. Participants

Twenty-four children (13 male, 11 female) approximately 34 months of age (range 33 months, 0 days to 34 months, 29 days) participated in this study. They did not differ in age from the children in experiment 1 [means of 34.2 vs 34.1 months, $t(46) = 0.71$, $p = 0.48$]. An additional ten children participated, but their data were excluded for excessive fussiness/exiting the test area/not attending ($n = 4$), being bilingual ($n = 2$), being outside the age range ($n = 1$), or experimenter error/failure to record the session ($n = 4$). An additional four participants were recruited in the event that additional data would be required in one of the stimulus orders, but their data (the last data collected in these orders) were not ultimately needed. Parents reported that their children had normal hearing and were not currently experiencing symptoms indicative of an ear infection.

Maternal education averaged 17.5 yr: three mothers had a doctoral degree, seven mothers had a master's degree, nine mothers had a four-year college degree, and one mother had an associate's degree. The average maternal education level did not differ from those of the children in experiment 1 [$t(46) = 1.08$, $p = 0.28$].

The final set of children were 67% Caucasian, 8% African American, 8% Asian, 13% Hispanic, and 4% mixed race. Parent-reported vocabulary on the Language Development Scale ranged from 120 to 310 words, and did not differ from the children in experiment 1 [$t(46) = 0.94$, $p = 0.35$]. Twelve children reported hearing another language in the home: one heard 1% Japanese, one heard 1% German, one heard 2% Arabic, one heard 1.5% French and 0.5% German, two heard 5% Mandarin, one heard 5% Yoruba, two heard 10% Spanish, one heard 15% Spanish, one heard 15% Hebrew, and one heard 20% Italian.

2. Materials, procedure, and coding

These were identical to experiment 1 except that the stimuli were vocoded with 16 channels rather than 8 channels.

B. Results and discussion

Here, children looked to the correct object 61.9% of the time (see the right side of Fig. 1). This is significantly different from chance [$t(23) = 4.14$, $p < 0.0005$, $d = 0.8642$], as well as significantly different from the results from experiment 1 [$t(46) = 2.97$, $p < 0.005$, $d = 0.7934$]. Thus, even though children failed to fast-map new words with the eight-channel vocoded speech in experiment 1, they were successful at fast-mapping new words when the signal was less degraded.

One concern is that children in the current study might have simply attended longer during the training phase than did the children hearing eight-channel vocoded speech, providing more opportunities to learn. An analysis of attention during the training showed that children in the 8-channel condition looked to the object on the screen an average of 89.2% of the time during the training phase; those in the 16-channel condition looked 91% of the time. This difference was not significant [$t(46) = 0.80$, $p = 0.427$], suggesting that there was no indication the 8-channel group was less attentive.

Looking at the performance of individual children, we see that children were generally clustered around chance performance in the eight-channel noise-vocoded condition. In the 16-channel condition, there was also a group of children with near-chance performance, but approximately half of the children showed clearly above-chance performance (performance above 0.6). This may be an indication of individual differences among children in their ability to fast-map from a degraded signal. However, performance was not statistically bimodal since normality was not violated. Moreover, performance among the children in experiment 2 did not correlate with vocabulary scores (and, indeed, went in the opposite direction; $r = -0.16$). This supports findings from a recent study examining NH school-age children's recognition of vocal emotion from vocoded stimuli (Tinnemore *et al.*, 2018); there, vocabulary was likewise not predictive of performance but nonverbal intelligence quotient (IQ) was. Thus, it is not clear what might be the underlying cause of performance differences among children in the current study, but general cognitive skills are a likely possibility.

IV. GENERAL DISCUSSION

In both studies, toddlers were first taught labels for two new objects. They were then presented with images of both objects at the same time and asked to look at one of the two objects. When the voice was presented in 16-channel noise-vocoded speech, the children were highly successful at looking at the correct object in the test phase. When the voice was presented in eight-channel noise-vocoded speech, however, they were not successful at looking at the correct object in the test phase. This suggests that the level of degradation imposed by eight-channel vocoding is sufficient to disrupt fast-mapping in young children.

Because the speech was presented as noise-vocoded both during learning and during the test, we cannot be certain whether children failed to learn the mapping in the first portion of the experiment or failed to recognize the speech in the second half, but there is good reason to believe that the learning phase was the problem. Most telling, the test phase of this study is essentially the same task as that presented in Newman and Chatterjee (2013) for already-known words. In that study, children slightly younger (27 months of age as compared to 34 months of age tested here) successfully recognized already-known words in eight-channel

noise-vocoded speech. Indeed, 20 out of 24 infants in that study looked longer to the named image, a highly consistent finding. There is no reason to expect that younger children would be successful in a recognition task while older children are not. Moreover, the prior task may actually have been more difficult in one way: in that study, children experienced a mismatch between their stored representations (which would have been based on full-spectrum speech) and the degraded signal presented at the test. In the present study, the words presented at the test had the same level of degradation as the stimuli that were initially learned. If a mismatch between signal and stored representation takes effort to overcome, one might have expected that the test portion of the current study would actually have been easier than the test portion of the prior work.

Another concern is that the difficulty may not have been in fast-mapping *per se* but perhaps was a result of the children not understanding the instructions during the test phase (e.g., “Find the”). This is unlikely for two reasons. First, even younger children succeeded in prior work using the same degradation level and testing task (Newman and Chatterjee, 2013). Second, children in preferential looking tasks do not appear to need sentence instructions to look at the appropriate object: they likewise look for isolated words even without a sentence context (Fernald and Hurtado, 2006; Tincoff and Jusczyk, 1999). Another possibility is that children may have been unable to hear the difference between the two target words, needoke and coopa (such that the failure was one of perception not learning). While we did not test children’s discrimination of these two words with eight-channel vocoding, the fact that children readily discriminated “car” from “ball” in the same task in the prior Newman and Chatterjee (2013) study leads us to believe that the results are not likely from a failure of discrimination. Still, we cannot entirely rule out that possibility.

The ability to learn from spectrally degraded speech is a critical skill in order for children with a CI to successfully process and learn their native language, and the ability to fast-map is an important contributor to this skill. While the degraded signal we use here is not identical to the signal CI learners actually hear, laboratory studies using vocoded signals (CI simulations) are often thought to represent the “best-case” scenario: how well listeners can interpret a speech signal when outside factors (such as nerve degeneration as a result of long-term lack of stimulation, etc.) are eliminated. If the current study is interpreted as a best case, the results would suggest that children with CIs may have difficulty fast-mapping new words through their implant, potentially slowing the process of word learning. Indeed, prior research has suggested that children with CIs do have difficulties with fast-mapping relative to their peers (Houston *et al.*, 2012; Tomblin *et al.*, 2007; Walker and McGregor, 2013).

Clearly children outside the laboratory do learn new words post-implantation. How, then, do they do so? One possibility is that learning from a degraded signal is simply slower than learning from full-spectrum speech: children

may need a greater number of exposures to a new word in order to successfully build a representation or may need to hear a word in multiple contexts. Despite having failed to show fast-mapping in the current task, they might have been successful had they been given more opportunities for slow, gradual word learning. To put it another way, a degraded signal may make learning more difficult but not prevent learning altogether. This would suggest that we might see successful learning in the laboratory if we provided children with more (and more varied) exposures. Comparing children’s performance learning new words with greater or fewer repetitions would be a fruitful direction for future research. More importantly, if children with CIs need more repetitions in order to learn new words, this would suggest that they would benefit from more exposure to word-learning situations, something that could be instantiated in clinical habilitation.

Alternatively, children with CIs may be more successful fast-mapping new words than were our current participants as a result of their substantial practice listening to a degraded signal in general. That is, early implanted CI children would have the advantage of developing their language system with degraded input, and this experience could allow them to overcome the limitations in fast-mapping from a degraded signal. In contrast, our NH children presumably did not have any experience listening to spectrally degraded speech. Whereas children apparently require very little experience in order to recognize words from a degraded signal (Newman and Chatterjee, 2013), the present results suggest they need more experience in order to learn from it. This necessity for further experience will add to the existing language delays caused by a lack of input pre-implantation and suggests that the actual “delay” in language learning experienced by children with a CI is likely to be longer than their age at implantation may suggest. Future work using noise-vocoded stimuli could attempt to assess how much experience is required before children are able to fast-map from a degraded signal.

Another possible explanation of the current results is that children may not need more experience to fast-map from a degraded signal compared to simply recognizing it but may instead need more cognitive resources. Listening to a degraded signal is likely to require more memory and attentional resources than does listening to full-spectrum speech, and this may reduce the amount of these resources available for storing information in long-term memory. There could be a compounding demand on children’s limited memory and attentional resources created by both the task of listening to a degraded signal and attempting to learn new words from it. Work from the speech recognition literature suggests that the pediatric CI population already requires more cognitive resources to process everyday speech than their NH peers (Grieco-Calub *et al.*, 2009). Perhaps the added cognitive requirements of learning new words simply overwhelm the limited resources available to young children. This would suggest that older children might well be successful at fast-mapping from eight-channel

noise-vocoded stimuli even without additional practice *per se*. This, too, points to a direction for future research. However, waiting until children are older (and letting the difficulty resolve itself) is unlikely to be a good strategy since children with CIs would fall further behind their peers. Moreover, children with actual CIs clearly do not fail to learn new words before the age of 34 months. Something, then, in their experience with a CI allows them to successfully bypass this apparent limit on fast-mapping from a degraded signal.

Another possible explanation of the current results ties in with young children's known difficulties with perceptual restoration (Newman, 2006). Adults and school-aged children regularly "fill-in" missing information from a speech signal (Bashford *et al.*, 1992; Bashford and Warren, 1987; Bashford *et al.*, 1996; Koroleva *et al.*, 1991; Layton, 1975; Newman, 2004; Samuel, 1981a,b, 1987; Samuel and Ressler, 1986; Warren, 1970; Warren and Obusek, 1971; Warren *et al.*, 1997; Warren and Sherman, 1974), allowing them to accurately perceive words even when information about some of the phonemes making up the word is absent. Toddlers, however, do not show this restoration ability (Newman, 2006), suggesting that they may have more difficulty with degraded signals such as the ones provided here.

Recent years have seen an increase in research investigating the impact of adverse conditions on speech recognition more broadly (for a review, see Mattys *et al.*, 2012). Much of this work has focused on situations in which the signal is accompanied by noise or another distractor. Many of these studies have shown relatively similar performance levels across word recognition and word learning. For example, children listening to speech in the presence of multi-talker babble can consistently recognize words at a 0 dB signal-to-noise ratio (SNR; that is, with the target speech at the same intensity as the masker; Newman, 2011) and showed some ability even beyond that (at a -5 dB SNR with the target speech less intense than the masker). Critically, they appear to be able to fast-map new words (in a task identical to the current one) at very similar noise levels (Dombroski and Newman, 2014). Thus, the presence of background babble appears to impact fast-mapping and word recognition quite similarly. In contrast, the current work suggests that spectral degradation may particularly impact fast-mapping relative to word recognition. A more thorough understanding of the impact of diverse adverse conditions, then, might need to focus differentially on their impact on learning as compared to recognition.

By these accounts, the current results suggest that children with CIs will either require additional exposure to new words in order to learn them (compared to NH peers), more well-developed cognitive skills, or more experience listening to a degraded signal in general before they can begin learning. In either case, the current study identifies an important hurdle faced by children with CIs. Word learning is likely to be an especial problem for children who are able to make use of fewer channels. Work with adults suggests that most CI users are not able to fully utilize the number of

channels provided by their implant (Friesen *et al.*, 2001), but this varies across individuals. Presumably, children likewise vary in this respect, and those who have access to fewer distinct channels may be more at risk for difficulties with fast-mapping and perhaps word learning more generally (and may need a greater amount of intervention).

Finally, the current work has implications well beyond those of children with CIs in that it highlights the profound difference between recognizing known words and learning new ones. As noted above, there are several possible reasons for this difference. Word learning and, specifically, fast-mapping may simply be more dependent on limited cognitive resources than word recognition and thus more susceptible to a poor-quality signal that places these cognitive resources under further demand. Alternatively, while recognition requires only that a word form be sufficiently identifiable as to distinguish it from other potential known words, fast-mapping may necessitate a more complete phonetic representation, something made more difficult by a degraded speech signal. Each of these is likely to impact children with NH in addition to those with CIs. Further study of the effects of noise-vocoded speech could help elucidate the underlying reasons why fast-mapping and word recognition differ.

V. CONCLUSION

The current study suggested that while children can fast-map new words from 16-channel degraded speech, they fail to do so with an 8-channel signal. This is despite the fact that they appear to have little difficulty recognizing already-learned words at this same degradation level. A signal with eight channels or less of spectral information is a reasonable estimate of the amount of signal degradation faced by the average CI listener, while high-performing patients gain benefit from more active channels (Chatterjee *et al.*, 2015; Friesen *et al.*, 2001). A recent study showed that adult CI listeners improved in a speech recognition task at the group level when the number of electrodes was increased from 8 to 20, but the improvement was incremental, and about half the CI participants did not show it (Croghan *et al.*, 2017). Given these considerations, we infer that spectro-temporal resolution in electric hearing is a limiting factor in pediatric CI users' acquisition of new words.

Yet, children with CIs do learn vocabulary in the real world, albeit at varying rates (Niparko *et al.*, 2010), and thus the signal must provide the means for them to do so. Future work is needed to explore how much experience with a degraded signal or particular novel words children require in order to overcome this difficulty in novel fast-mapping.

ACKNOWLEDGMENTS

The authors thank George Hollich for the Supercoder program, and Arielle Abrams, Alyssa Ambielli, Sarah Aylor, Jackie Berges, Ariella Bloomfield, Tiara Booth, Kelly Cavanaugh, Rachel Childress, Jaya Chinnaya, Cathy Eaton, Lyana Frantz, Katherine Gagan, Gabrielle Giangrasso, Abby Goron, Laura Goudreau, Devin Heit,

Chantal Hoff, Koroush Kalachi, Greer Kauffman, Caroline Kettl, Penina Kozlovsky, Aliza Layman, Hannah Lebovics, Chandler Littleton, Laura Miller, Ariya Mobaraki, Catie Penny, Emma Peterson, Andrea Picciotto, Carly Pontell, Mariah Pranger, Kelly Puyear, Hallie Saffair, Asim Shafique, Rebecca Sherman, Naomi Silverman, Emily Slonecker, Sydney Smith, Lydia Sonenklar, Lauren Steedman, Anna Stone, Ashley Thomas, Nicole Tobin, Allison Urbanus, Krista Voelmle, Natalie Walter, Catherine Wilson, Kimmie Wilson, Rebecca Wolf, Rami Yanes, and Erica Younkin for assistance in stimulus recording, scheduling or testing participants, and coding looking time performances. This work was supported by the National Institutes of Health (NIH) Grant No. R01 HD081127 to the University of Maryland.

- Bashford, J. A., Riener, K. R., and Warren, R. M. (1992). "Increasing the intelligibility of speech through multiple phonemic restorations," *Percept. Psychophys.* **51**(3), 211–217.
- Bashford, J. A., and Warren, R. M. (1987). "Multiple phonemic restorations follow the rules for auditory induction," *Percept. Psychophys.* **42**(2), 114–121.
- Bashford, J. A., Warren, R. M., and Brown, C. A. (1996). "Use of speech-modulated noise adds strong 'bottom-up' cues for phonemic restoration," *Percept. Psychophys.* **58**(3), 342–350.
- Baskent, D., and Shannon, R. V. (2003). "Speech recognition under conditions of frequency-place compression and expansion," *J. Acoust. Soc. Am.* **113**(4), 2064–2076.
- Baskent, D., and Shannon, R. V. (2006). "Frequency transposition around dead regions simulated with a noiseband vocoder," *J. Acoust. Soc. Am.* **119**(2), 1156–1163.
- Baskent, D., and Shannon, R. V. (2007). "Combined effects of frequency compression-expansion and shift on speech recognition," *Ear Hear.* **28**(3), 277–289.
- Berg, K. A., Noble, J. H., Dawant, B. M., Dwyer, R. T., Labadie, R. F., and Gifford, R. H. (2019). "Speech recognition as a function of the number of channels in perimodiolar electrode recipients," *J. Acoust. Soc. Am.* **145**(3), 1556–1564.
- Bion, R. A. H., Borovsky, A., and Fernald, A. (2013). "Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months," *Cognition* **126**(1), 39–53.
- Bloom, P. (2000). *How Children Learn the Meanings of Words* (MIT Press, Cambridge, MA).
- Boons, T., Brokx, J. P. L., Dhooge, I., Frijns, J. H., Peeraer, L., Vermeulen, A., Wouters, J., and van Wieringen, A. (2012). "Predictors of spoken language development following pediatric cochlear implantation," *Ear Hear.* **33**(5), 617–639.
- Carey, S. (2010). "Beyond fast mapping," *Lang. Learn. Dev.* **6**(3), 184–205.
- Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., Kulkarni, A. M., and Christensen, J. A. (2015). "Voice emotion recognition by cochlear-implanted children and their normally-hearing peers," *Hear. Res.* **322**, 151–162.
- Croghan, N. B. H., Duran, S. I., and Smith, Z. M. (2017). "Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility," *J. Acoust. Soc. Am.* **142**(6), EL537–EL542.
- Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., and McGarrigan, C. (2005). "Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences," *J. Exp. Psychol. Gen.* **134**(2), 222–241.
- Dombroski, J., and Newman, R. S. (2014). "Toddlers' ability to map the meaning of new words in multi-talker environments," *J. Acoust. Soc. Am.* **136**(5), 2807–2815.
- Dorman, M. F., Loizou, P. C., Kemp, L. L., and Kirk, K. I. (2000). "Word recognition by children listening to speech processed into a small number of channels: Data from normal-hearing children and children with cochlear implants," *Ear Hear.* **21**(6), 590–596.
- Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., and Boothroyd, A. (2000). "Speech recognition with reduced spectral cues as a function of age," *J. Acoust. Soc. Am.* **107**(5), 2704–2710.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., and Pethick, S. J. (1994). "Variability in early communicative development," *Monogr. Soc. Res. Child Dev.* **59**(5), Serial 242, 1–173.
- Fernald, A., and Hurtado, N. (2006). "Names in frames: Infants interpret words in sentence frames faster than words in isolation," *Dev. Sci.* **9**(3), F33–F40.
- Fishman, K. E., Shannon, R. V., and Slattery, W. H. (1997). "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," *J. Speech, Lang. Hear. Res.* **40**(5), 1201–1215.
- Friese, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *J. Acoust. Soc. Am.* **110**(2), 1150–1163.
- Fu, Q. J., and Shannon, R. V. (1998). "Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners," *J. Acoust. Soc. Am.* **104**(5), 2570–2577.
- Fu, Q. J., and Shannon, R. V. (1999). "Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing," *J. Acoust. Soc. Am.* **105**(3), 1889–1900.
- Fu, Q. J., Shannon, R. V., and Wang, X. S. (1998). "Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing," *J. Acoust. Soc. Am.* **104**(6), 3586–3596.
- Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., and Gordon, L. (1987). "The eyes have it: Lexical and syntactic comprehension in a new paradigm," *J. Child Lang.* **14**, 23–45.
- Golinkoff, R. M., Ma, W., Song, L., and Hirsh-Pasek, K. (2013). "Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned?," *Perspect. Psychol. Sci.* **8**(3), 316–339.
- Grieco-Calub, T. M., Saffran, J. R., and Litovsky, R. Y. (2009). "Spoken word recognition in toddlers who use cochlear implants," *J. Speech, Lang. Hear. Res.* **52**, 1390–1400.
- Hervais-Adelman, A., Davis, M. H., Johnsrude, I. S., and Carlyon, R. P. (2008). "Perceptual learning of noise vocoded words: Effects of feedback and lexicality," *J. Exp. Psychol. Hum. Percept. Perform.* **34**(2), 460–474.
- Hollich, G. (2005). "Supercoder: A program for coding preferential looking (version 1.5) [computer program]," Purdue University, West Lafayette, IN, available at <http://hincapie.psych.purdue.edu/Splitscreen/>.
- Hollich, G. (2006). "Combining techniques to reveal emergent effects in infants' segmentation, word learning, and grammar," *Lang. Speech* **49**(1), 3–19.
- Horst, J. S., and Samuelson, L. K. (2008). "Fast mapping but poor retention by 24-month-old infants," *Infancy* **13**(2), 128–157.
- Houston, D. M., Stewart, J., Moberly, A., Hollich, G., and Miyamoto, R. T. (2012). "Word learning in deaf children with cochlear implants: Effects of early auditory experience," *Dev. Sci.* **15**(3), 448–461.
- Ihlefeld, A., Deeks, J. M., Axon, P. R., and Carlyon, R. P. (2010). "Simulations of cochlear-implant speech perception in modulated and unmodulated noise," *J. Acoust. Soc. Am.* **128**(2), 870–880.
- Jahn, K. N., DiNino, M., and Arenberg, J. G. (2019). "Reducing simulated channel interaction reveals differences in phoneme identification between children and adults with normal hearing," *Ear Hear.* **40**(2), 295–311.
- Koroleva, I. V., Kashina, I. A., Sakhnovskaya, O. S., and Shurgaya, G. G. (1991). "Perceptual restoration of a missing phoneme: New data on speech perception in children," *Sens. Syst.* **5**(3), 191–199, available at <https://psycnet.apa.org/record/1992-34543-001>.
- Kucker, S. C., McMurray, B., and Samuelson, L. K. (2015). "Slowing down fast mapping: Redefining the dynamics of word learning," *Child Dev. Perspect.* **9**(2), 74–78.
- Layton, B. (1975). "Differential effects of two nonspeech sounds on phonemic restoration," *Bull. Psychon. Soc.* **6**(5), 487–490.
- Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scot, S. K. (2012). "Speech recognition in adverse conditions: A review," *Lang. Cognit. Processes* **27**, 953–978.
- McMurray, B. (2007). "Defusing the childhood vocabulary explosion," *Science* **317**, 631.
- McMurray, B., Horst, J. S., and Samuelson, L. K. (2012). "Word learning emerges from the interaction of online referent selection and slow associative learning," *Psychol. Rev.* **119**(4), 831–877.

- National Institute on Deafness and Other Communication Disorders. (2016). "NIDCD fact sheet: Cochlear implants," NIH Publication No. 00-4798, available at <https://www.nidcd.nih.gov/sites/default/files/Documents/health/hearing/FactsheetCochlearImplants.pdf> (Last viewed 4/15/20).
- Newman, R. S. (2004). "Perceptual restoration in children versus adults," *Appl. Psycholing.* **25**, 481–493.
- Newman, R. S. (2006). "Perceptual restoration in toddlers," *Percept. Psychophys.* **68**, 625–642.
- Newman, R. S. (2011). "2-year-olds' speech understanding in multi-talker environments," *Infancy* **16**(5), 447–470.
- Newman, R. S., and Chatterjee, M. (2013). "Toddlers' recognition of noise-vocoded speech," *J. Acoust. Soc. Am.* **133**(1), 483–494.
- Newman, R. S., Chatterjee, M., Morini, G., and Remez, R. E. (2015). "Toddlers' comprehension of degraded signals: Noise-vocoded versus sine-wave analogs," *J. Acoust. Soc. Am.* **138**(3), EL311–EL317.
- Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N. Y., Quittner, A. L., Fink, N. E., and the CDaCI Investigative Team. (2010). "Spoken language development in children following cochlear implantation," *J. Am. Med. Soc.* **303**(15), 1498–1506.
- Nittrouer, S., and Lowenstein, J. H. (2010). "Learning to perceptually organize speech signals in native fashion," *J. Acoust. Soc. Am.* **127**(3), 1624–1635.
- Nittrouer, S., Lowenstein, J. H., and Packer, R. R. (2009). "Children discover the spectral skeletons in their native language before the amplitude envelopes," *J. Exp. Psychol. Hum. Percept. Perform.* **35**, 1245–1253.
- Rescorla, L. (1989). "The Language Development Survey: A screening tool for delayed language in toddlers," *J. Speech Hear. Disord.* **54**(4), 587–599.
- Samuel, A. G. (1981a). "Phonemic restoration: Insights from a new methodology," *J. Exp. Psychol. Gen.* **110**, 474–494.
- Samuel, A. G. (1981b). "The role of bottom-up confirmation in the phonemic restoration illusion," *J. Exp. Psychol. Hum. Percept. Perform.* **7**, 1124–1131.
- Samuel, A. G. (1987). "Lexical uniqueness effects on phonemic restoration," *J. Mem. Lang.* **26**, 36–56.
- Samuel, A. G., and Ressler, W. H. (1986). "Attention within auditory word perception: Insights from the phonemic restoration illusion," *J. Exp. Psychol. Hum. Percept. Perform.* **12**(1), 70–79.
- Schwartz-Leyzac, K. C., Zwolan, T. A., and Pfingst, B. E. (2017). "Effects of electrode deactivation on speech recognition in multichannel cochlear implant recipients," *Cochlear Implants Int.* **18**(6), 324–334.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**(5234), 303–304.
- Shannon, R. V., Zeng, F. G., and Wygonski, J. (1998). "Speech recognition with altered spectral distribution of envelope cues," *J. Acoust. Soc. Am.* **104**(4), 2467–2476.
- Sheldon, S., Pichora-Fuller, M. K., and Schneider, B. A. (2008). "Priming and sentence context support listening to noise-vocoded speech by younger and older adults," *J. Acoust. Soc. Am.* **123**(1), 489–499.
- Swingle, D. (2010). "Fast mapping and slow mapping in children's word learning," *Lang. Learn. Dev.* **6**, 179–183.
- Tincoff, R., and Jusczyk, P. W. (1999). "Some beginnings of word comprehension in 6-month-olds," *Psychol. Sci.* **10**(2), 172–175.
- Tinnemore, A. R., Zion, D. J., Kulkarni, A. M., and Chatterjee, M. (2018). "Children's recognition of emotional prosody in spectrally degraded speech is predicted by their age and cognitive status," *Ear Hear.* **39**(5), 874–880.
- Tomblin, J. B., Barker, B. A., and Hubbs, S. (2007). "Developmental constraints on language development in children with cochlear implants," *Int. J. Audiol.* **46**, 512–523.
- Walker, E. A., and McGregor, K. K. (2013). "Word learning processes in children with cochlear implants," *J. Speech, Lang. Hear. Res.* **56**(2), 375–387.
- Warren, R. M. (1970). "Perceptual restoration of missing speech sounds," *Science* **167**, 392–393.
- Warren, R. M., and Obusek, C. J. (1971). "Speech perception and phonemic restorations," *Percept. Psychophys.* **9**(3-B), 358–362.
- Warren, R. M., Reiner Hainsworth, K., Brubaker, B. S., Bashford, J. A., Jr., and Healy, E. W. (1997). "Spectral restoration of speech: Intelligibility is increased by inserting noise in spectral gaps," *Percept. Psychophys.* **59**(2), 275–283.
- Warren, R. M., and Sherman, G. L. (1974). "Phonemic restorations based on subsequent context," *Percept. Psychophys.* **16**(1), 150–156.
- Werker, J. F., and Curtin, S. (2005). "PRIMIR: A developmental framework of infant speech processing," *Lang. Learn. Dev.* **1**(2), 197–234.
- Zekveld, A. A., and Kramer, S. E. (2014). "Cognitive processing load across a wide range of listening conditions: Insights from pupillometry," *Psychophysiology* **51**(3), 277–284.